UMLS® KNOWLEDGE SOURCES

**November Release
2005AC
DOCUMENTATION**

Section 0
Preface

0.1 Purpose of This Documentation

The Unified Medical Language System® (UMLS®) Documentation describes the UMLS Knowledge Sources and related tools that are produced and distributed by the National Library of Medicine, a part of the National Institutes of Health in the U.S. Department of Health and Human Services. This documentation explains:

- the purpose, content, and file structures of the current versions of the three UMLS Knowledge Sources - the Metathesaurus®, the Semantic Network, and the SPECIALIST Lexicon;
- why and how to use associated UMLS programs, including:
    (a) MetamorphoSys, the install program for the three UMLS Knowledge Sources; it also provides a variety of output options, including Rich Release and Original Release Formats, and a choice of preferred character set, to produce custom subsets of the Metathesaurus.
    (b) the Lexical Programs, which help to deal with inflectional variation (e.g., treat, treats, treating, treatment) in the English language; to convert American English to British English and vice versa, and to map text to concepts in the Metathesaurus;

- how to access the UMLS resources via the UMLS Knowledge Source Server via (a) download, (b) an application programming interface (API), or (c) a Web browser.
- the DVD-ROM distribution format for the UMLS Knowledge Sources and associated UMLS programs, which is available on request.

0.2 Release Schedule
This documentation is updated for each release of the UMLS. It may be updated between releases if errors or omissions are found. The current document refers to the 2005AC version which was released in November 2005. The next UMLS version will be 2006AA, scheduled to be released in January 2006.

0.3 Audience
This documentation and the UMLS resources it describes are intended for system developers, informatics researchers, librarians, and other information professionals. The documentation assumes that you are familiar with database concepts and the Internet. If you intend to use the UMLS Knowledge Sources in software applications, it assumes that you have experience with building and using complex databases. If you intend to use any of the UMLS programs, it assumes basic familiarity with Java.

Neither the UMLS resources nor this documentation are intended for "end" users, such as individual health professionals or members of the general public - unless they are also software developers.

0.4 How to Use This Documentation

0.4.0 If you are an experienced UMLS user
If you have done substantive work with preceding versions of the UMLS resources go directly to Section 0.5, which describes what has been changed for this new version - in the documentation and in the UMLS resources themselves. Section 0.5 will point you to the parts of the documentation that describe any changes to data files, content or format introduced in this release.

0.4.1 If you are a novice UMLS user
If you are new to the UMLS, Section 0.5 won't be particularly relevant (or understandable!), but you should read the rest of Section 0 and all of Section 1 before moving on to other parts of the documentation. Please read this brief overview which explains what you will find in each section of the documentation.

Section 1. Introduction to the UMLS

The section explains the purpose of the UMLS, explains the conditions under which you may use the different UMLS components and how these relate to Open Access/Open Source principles, provides a brief description of each of the UMLS components and the relationships between them, describes how to get a feel for what is in each of the UMLS Knowledge Sources, and provides a list of additional reference materials about the UMLS.

Section 2. Metathesaurus

This section describes the content and structure of the Metathesaurus, a very large concept-oriented database that incorporates many different biomedical and health-related vocabularies, classifications, and coding systems. The Metathesaurus provides a consistent categorization of these concepts by assigning basic semantic types; and makes all information from these terminologies accessible in common, fully-specified file formats. The Metathesaurus contains coding systems and vocabularies designated as U.S. standards under the Administrative Simplification provisions of the Health Insurance Portability and Accountability Act of 1996 (HIPAA) and as target U.S. government-wide standards by the Consolidated Health Informatics eGov initiative.

Section 3. Semantic Network

This section describes the content and structure of the Semantic Network, a small database that includes information about the set of basic semantic types, or categories, to which Metathesaurus concepts may be assigned. The Semantic Network defines the relationships that may hold between these semantic types, as well as broad groupings of semantic types, such as all types that denote disorders (Disease or Syndrome, Acquired Abnormality, Neoplastic Process, etc.).

Section 4. SPECIALIST Lexicon and Lexical Programs

This section describes the content and structure of (1) the SPECIALIST lexicon, a database of syntactic, morphological, and orthographic information for commonly occurring English language words and biomedical vocabulary that is useful for natural language processing applications and (2) lexical programs, which are designed to assist in detecting and abstracting away from the inflectional, case, and word order variations encountered in natural language. One of these, MMTx (Metamap Transfer), is specifically designed to map arbitrary terms to concepts in the Metathesaurus or, equivalently, to discover Metathesaurus concepts within free text.

Section 5. UMLS Knowledge Source Server

This section describes how to access the UMLS resources from the UMLS Knowledge Source Server via download, application programmer interface, and interactive Web browser.

Section 6. MetamorphoSys: the UMLS install and customization program

This section describes MetamorphoSys, the install program for all the UMLS Knowledge Sources and the customization program for the Metathesaurus. UMLS users *must* use MetamorphoSys to install the Knowledge Sources. MetamophoSys allows users to output data in either the 7-bit ASCII (the default) or Unicode UTF-8 character set. MetamorphoSys provides two file format options (Rich Release, or Original Release Format) for the Metathesaurus, and provides a number of other options to customize it in specific applications.

Section 7. UMLS DVD

This section gives technical specifications for the UMLS DVD-ROM, an alternative method for distribution of UMLS content, which is available on request to UMLS licensees.

Appendices

Appendix A includes the text of the License Agreement for Use of the UMLS Metathesaurus and Appendix.

Appendix B includes additional information about the Metathesaurus data elements and the source vocabularies in the current Metathesaurus.

0.5 What's New for This Version of the UMLS

0.5.0 Metathesaurus

The Metathesaurus now contains more than 1 million concepts and 5 million unique concept names from more than 100 different source vocabularies. Twenty-three source vocabularies were updated; including HL7 version 2, HL7 version 3, SNOMED CT, GO, and MESH. Two new sources, Common Terminology Criteria for Adverse Events v3.0 and United States Pharmacopeia were added.

0.5.1 MetamorphoSys

Welcome Screen

Both the RRF Browser and MRCXT Builder have been integrated into MetamorphoSys, these tools can be accessed directly from the welcome screen. Users who complete a RRF subset will be promted to open the RRF Broweser at the end of the subsetting process.

RxNorm Filter

The RxNorm default subset has been removed. Users can produce this subset with the RxNorm filter. This filter is accessed from the file menu. For more information see section 6.6.1.5.

RRF Browser

Users can now view metadata in the Raw View. Users can set the number of records displayed per page. Once set this feature allows users to navigate from page to page using "prev" and "next" links. In addition users who complete a RRF subset will be promted to open the RRF Broweser at the end of the subsetting process. Information on the RRF Browser will be found at: http://www.nlm.nih.gov/research/umls/rrf_help.html.

0.5.2 Coming Changes

Important notice: Beginning with the 2006AA Release the file MRCXT will no longer be distributed. The information contained in this large and unwieldy file of pre-computed contexts is better represented in MRHIER.RRF. Users can create MRCXT files for their own customized subsets using the new MRCXT Builder (above). As a result the 2005AA UMLS Release will contain smaller file sizes and run faster than previous releases. Sources scheduled to be updated include CPT, SNOMED Spanish, & MEDRA 8.0.

Table of Contents

**Section 1
Introduction to the UMLS**

1.1 Purpose of the UMLS

The purpose of NLM's Unified Medical Language System (UMLS) is to facilitate the development of computer systems that behave as if they "understand" the meaning of the language of biomedicine and health. To that end, NLM produces and distributes the UMLS Knowledge Sources (databases) and associated software tools (programs) for use by system developers in building or enhancing electronic information systems that create, process, retrieve, integrate, and/or aggregate biomedical and health data and information, as well as in informatics research. By design, the UMLS Knowledge Sources are multi-purpose. They are not optimized for particular applications, but can be applied in systems that perform a range of functions involving one or more types of information, e.g., patient records, scientific literature, guidelines, public health data. The associated UMLS software tools assist developers in customizing or using the UMLS Knowledge Sources for particular purposes. The lexical tools work more effectively in combination with the UMLS Knowledge Sources, but can also be used independently.

1.2 Conditions of Use of the UMLS

All UMLS Knowledge Sources and associated software tools are free of charge to U.S. and international users.

The Semantic Network, the SPECIALIST Lexicon, and associated lexical tools are freely accessible on the Internet under open *terms*, which include appropriate acknowledgment for their use. View the terms and conditions for use of the Semantic Network and of the SPECIALIST Lexicon and Lexical Tools.

To use the Metathesaurus, you must establish a license agreement. This is because the Metathesaurus includes vocabulary content produced by many different copyright holders as well as the substantial content produced by NLM.

Do NOT let the license requirement discourage you from using the Metathesaurus. Setting up the license agreement is quick and easy and is done via the Web. Once the license agreement is in place, much of the content of the Metathesaurus may be used under very open conditions. Your pre-existing licenses for content with use restrictions, e.g., CPT, MedDRA, or NIC, will cover your use of that content as distributed within the Metathesaurus. Some vocabulary producers who require you to request permission for production uses of their content will generally grant permission free of charge.

The complete text of the License Agreement for Use of the UMLS Metathesaurus appears in Appendix A of this documentation.

1.3 Brief Description of the UMLS Knowledge Sources and Associated Tools

There are three UMLS Knowledge Sources: the Metathesaurus, the Semantic Network, and the SPECIALIST Lexicon. They are distributed with several tools (programs) that facilitate their use, including the MetamorphoSys install and customization program.

1.3.1 Metathesaurus

The Metathesaurus is a very large, multi-purpose, and multi-lingual vocabulary database that contains information about biomedical and health-related concepts, their various names, and the relationships among them. It is built from the electronic versions of many different thesauri, classifications, code sets, and lists of controlled terms used in patient care, health services billing, public health statistics, indexing and cataloging biomedical literature, and/or basic, clinical, and health services research. In this documentation, these are referred to as the "source vocabularies" of the Metathesaurus. In the Metathesaurus, all the source vocabularies are available in a single, fully-specified database format.

A complete list of the source vocabularies present in this version of the Metathesaurus appears in Appendix B.4 to this documentation. The list indicates which coding systems and vocabularies are designated as U.S. standards for administrative health transactions in accordance with HIPAA or as target U.S. government-wide clinical standards selected by the Consolidated Health Informatics eGov initiative.

The Metathesaurus is organized by concept or meaning. In essence, its purpose is to link alternative names and views of the same concept together and to identify useful relationships between different concepts. All concepts in the Metathesaurus are assigned to at least one semantic type from the Semantic Network (1.3.2). This provides consistent categorization of all concepts in the Metathesaurus at the relatively general level represented in the Semantic Network. Many of the words and multi-word terms that appear in concept names or strings in the Metathesaurus also appear in the SPECIALIST Lexicon (1.3.3.1). The lexical tools (1.3.3.2) are used to generate the word, normalized word, and normalized string indexes to the Metathesaurus. MetamorphoSys (1.3.5) must be used to install all the UMLS Knowledge Sources and is the recommended software tool for customizing the Metathesaurus.

The Metathesaurus *must* be customized to be used effectively.

A complete description of the Metathesaurus and its file structure appears in Section 2 of this documentation.

1.3.2 Semantic Network

The purpose of the Semantic Network is to provide a consistent categorization of all concepts represented in the UMLS Metathesaurus and to provide a set of useful relationships between these concepts. All information about specific concepts is found in the Metathesaurus; the Network provides information about the set of basic semantic types, or categories, which may be assigned to these concepts, and it defines the set of relationships that may hold between the semantic types. The current release of the Semantic Network contains 135 semantic types and 54 relationships. The Semantic Network serves as an authority for the semantic types that are assigned to concepts in the Metathesaurus. The Network defines these types, both with textual descriptions and by means of the information inherent in its hierarchies.

The Semantic Types are the nodes in the Network, and the Semantic Relations between them are the links. There are major groupings of semantic types for organisms, anatomical structures, biologic function, chemicals, events, physical objects, and concepts or ideas. The current scope of the UMLS semantic types is quite broad, allowing for the semantic categorization of a wide range of terminology in multiple domains.

A complete description of the Semantic Network and its file structure appears in Section 3 of this documentation.

1.3.3 SPECIALIST Lexicon and Lexical Programs

The SPECIALIST Lexicon is intended to be a general English lexicon that includes many biomedical terms. Coverage includes both commonly occurring English words and biomedical vocabulary. The lexicon entry for each word or term records the syntactic, morphological, and orthographic information needed by the SPECIALIST Natural Language Processing

System.

The lexical programs or tools are designed to address the high degree of variability in natural language words and terms. Words often have several inflected forms which would properly be considered instances of the same word. The verb "treat", for example, has three inflectional variants: "treats" the third person singular present tense form, "treated" the past and past participle form, and "treating" the present participle form. Multi-word terms in the Metathesaurus and other controlled vocabularies may have word order variants in addition to their inflectional and alphabetic case variants. The lexical tools allow the user to abstract away from several types of variation, including British English/American English spelling variation and character set variations.

A complete description of the SPECIALIST Lexicon, its file structure, and the lexical programs appears in Section 4 of this documentation.

1.3.4 UMLS Knowledge Source Server
The UMLS Knowledge Source Server (UMLSKS) is a set of Web-based interactive tools and a programmer interface to allow users and developers access to the UMLS Knowledge Sources, including the vocabularies within the Metathesaurus. It also contains the download site for the UMLS data files. The UMLSKS is a useful starting point for gaining an understanding of the content of the UMLS resources. Because it contains the complete Metathesaurus files, access to UMLSKS is restricted to registered users who have signed the License Agreement for Use of the UMLS Metathesaurus.

A complete description of the UMLS Knowledge Source server and its capabilities appears in Section 5 of this documentation.

1.3.5 MetamorphoSys - the UMLS install and customization program
MetamorphoSys is a cross-platform Java application that must be used if the UMLS Knowledge Sources (Metathesaurus, Semantic Network, and SPECIALIST Lexicon) are installed locally. MetamorphoSys also supports the creation and refinement of customized subsets of the Metathesaurus. In general, the Metathesaurus must be customized to be used effectively in specific applications.

MetamorphoSys guides you first through the installation of one or more UMLS Knowledge Sources, and then through customization of the Metathesaurus for local use. A variety of options are available, such as the inclusion or exclusion of specific source vocabularies, languages, and term types, specification of output character set (7-bit ASCII or Unicode UTF-8) and output format (Rich Release Format or Original Release Format) for the Metathesaurus files.

A complete description of MetamorphoSys appears in Section 6 of this documentation.

1.4 Getting Started
The UMLS resources are powerful - and unusual - tools intended for use by system developers. Here are a few suggestions about how to start building your understanding of UMLS features and capabilities and their potential for enhancing your applications.

Scan the entire UMLS documentation to get a sense of the range of resources available.

If the Metathesaurus interests you, take time to read Sections 2.1-2.6 of the documentation. The background there will make it easier to understand the actual file descriptions in Section 2.7.

Use the Web registration system to execute the free License Agreement for Use of the UMLS Metathesaurus. A license agreement is required because the Metathesaurus contains vocabularies produced by many different copyright holders. You are able to use much of the content of the Metathesaurus with minimal restriction, but you may need to obtain additional licenses from individual vocabulary producers if you wish to use certain vocabularies contained in the Metathesaurus. The various restriction levels are explained in the UMLS license agreement and its Appendix.

Once you have executed the current License agreement, use the UMLS Knowledge Source Server for initial browsing and exploration of the contents of the Metathesaurus, Semantic Network, and SPECIALIST Lexicon and of additional special

resources available useful to application developers.

If you require local copies of the UMLS files, use the MetamorphoSys install and customization program described in Section 6 to produce them. You may find it useful to experiment with various options to produce customized subsets. MetamorphoSys comes on the UMLS DVD and is available for download with the UMLS data files from the UMLS Knowledge Source Server.

1.5 Sources of Additional Information about the UMLS
In addition to providing links to the UMLS documentation and to the UMLS Knowledge Source Server, NLM's UMLS website at: http://umlsinfo.nlm.nih.gov provides pointers to: current fact sheets on the UMLS Knowledge Sources and Knowledge Source Server; FAQs; training materials; and information about NLM applications and research projects that make use of the UMLS. Articles on the UMLS project and the use of UMLS resources can be retrieved from MEDLINE/ PubMed. Click here to obtain a current search. A comprehensive 1986-1996 bibliography on the UMLS project covering additional papers not indexed for MEDLINE/PubMed is also available.

UMLS users are strongly encouraged to subscribe to the umls users listserve. NLM uses this mechanism to seek advice from UMLS users and to distribute news about upcoming UMLS developments. UMLS users use the listserve to share experiences or obtain advice about using the UMLS resources.

To subscribe, send an email to: listserv@list.nih.gov containing the following message: SUBSCRIBE|SIGNUP UMLSUSERS-L <your full name>
To post a message to the list AFTER subscribing, send email to: UMLSUSERS-L@list.nih.gov
To access subscription information and list archives, go to: UMLSUSERS-L Listserv Webpage

An alternative list, UMLS-ANNOUNCES-L, exists for users who wish to receive only official announcements about UMLS products and services, including new releases, new features, and problem/fix messages.

To subscribe, send an email to: listserv@list.nih.gov containing the following message: SUBSCRIBE|SIGNUP UMLS-ANNOUNCES-L <your full name>
To access subscription information and list archives, go to: UMLS-ANNOUCES-L Listserv Webpage

Specific questions about the UMLS can be addressed to custserv@nlm.nih.gov or, for telephone inquiries, to 1-888-FINDNLM (1-888-346-3656).

Table of Contents

Section 2
Metathesaurus

2.0 Introduction

The Metathesaurus is a very large, multi-purpose, and multi-lingual vocabulary database that contains information about biomedical and health related concepts, their various names, and the relationships among them. Designed for use by system developers, the Metathesaurus is built from the electronic versions of many different thesauri, classifications, code sets, and lists of controlled terms used in patient care, health services billing, public health statistics, indexing and cataloging biomedical literature, and/or basic, clinical, and health services research. These are referred to as the "source vocabularies" of the Metathesaurus. The term Metathesaurus draws on Webster's Dictionary third definition for the prefix "Meta," i.e., "more comprehensive, transcending." In a sense, the Metathesaurus transcends the specific thesauri, vocabularies, and classifications it encompasses.

The Metathesaurus is organized by concept or meaning. In essence, its purpose is to link alternative names and views of the same concept together and to identify useful relationships between different concepts.

The Metathesaurus is linked to other UMLS Knowledge Sources. All concepts in the Metathesaurus are assigned to at least one semantic type from the Semantic Network (Section 3.0). This provides consistent categorization of all concepts in the Metathesaurus at the relatively general level represented in the Semantic Network. Many of the words and multi-word terms that appear in concept names or strings in the Metathesaurus also appear in the SPECIALIST lexicon (Section 4). The lexical tools (Section 4) are used to generate the word, normalized word, and normalized string indexes to the Metathesaurus. MetamorphoSys (Section 6) is the software tool for customizing the Metathesaurus for specific purposes. It is also the install program for all of the UMLS resources.

2.0.1 Scope of the Metathesaurus

The scope of the Metathesaurus is determined by the combined scope of its source vocabularies. Many relationships (primarily synonymous), concept attributes, and some concept names are added by the NLM during Metathesaurus creation and maintenance, but essentially all the concepts themselves come from one or more of the source vocabularies. With very few exceptions, if none of the source vocabularies contains a concept, that concept will not appear in the Metathesaurus.

2.0.2 Preservation of Content and Meaning from Source Vocabularies

The Metathesaurus reflects and preserves the meanings, concept names, and relationships from its source vocabularies. When two different source vocabularies use the same name for differing concepts, the Metathesaurus represents both of the meanings and indicates which meaning is present in which source vocabulary. When the same concept appears in different hierarchical contexts in different source vocabularies, the Metathesaurus includes all the hierarchies. When conflicting relationships between two concepts appear in different source vocabularies, both views are included in the Metathesaurus. Although specific concept names or relationships from some source vocabularies may be idiosyncratic and lack face validity, they are still included in the Metathesaurus.

In other words, the Metathesaurus does not represent a comprehensive NLM-authored ontology of biomedicine or a single consistent view of the world (except at the high level of the semantic types assigned to all its concepts). The Metathesaurus preserves the many views of the world present in its source vocabularies because these different views may be useful for different tasks.

Although it preserves all the meanings and content in its source vocabularies, the Metathesaurus stores this information in a single common format. The native format of each vocabulary is carefully studied and then "inverted" into the common Metathesaurus format. For some vocabularies, this involves representing implied information in a more explicit format. To give an example, if a source vocabulary stores its preferred concept name as the first occurrence in a list of alternative concept names, that first name is explicitly tagged as the preferred name for that source in the Metathesaurus.

2.0.3 Need to Customize the Metathesaurus

Because it is a multi-purpose resource that includes concepts and terms from many different source vocabularies developed for very different purposes**, the Metathesaurus *must* be customized for effective use in most specific applications.** Your decisions about what to include in your customized subset(s) of the Metathesaurus will have a significant effect on its utility in your systems. Vocabulary sources that are essential for some purposes, e.g., LOINC for standard exchange of laboratory data, may be detrimental for others, such as natural language processing. It can also be important to exclude a subset of the concept names found in a vocabulary source that is otherwise useful, e.g., non-standard abbreviations or shortened forms that lack face validity or produce spurious results in natural language processing.

The Metathesaurus contains source vocabularies produced by many different copyright holders. The majority of the content of the Metathesaurus is available for use under the basic (and quite open) terms described in sections 1-11 and 13-16 of the Metathesaurus license. However, some vocabulary producers place additional restrictions on the use of their content as distributed within the Metathesaurus. The various levels of additional restrictions are described in Section 12 of the license. The level that applies to individual vocabularies is recorded in the Appendix to the license in Appendix B.4 to this documentation, and in the MetamorphoSys install and customization program (Section 6.0). If a UMLS user already has a separate license for use of one of the source vocabularies, the user's existing license also applies to that source as distributed within the Metathesaurus. In some cases, UMLS users may have to request permission or negotiate a separate license with a

vocabulary producer in order to use that vocabulary in a production system. There may be a charge associated with these separate permissions or license agreements.

The Metathesaurus is designed to facilitate customization. All information in the Metathesaurus is labeled as to its source (s), so it is possible to determine which concept names, attributes, and relationships come from which source vocabularies and which attributes and relationships were added during Metathesaurus construction. The labels allow UMLS users to subset the Metathesaurus by excluding information from specific source vocabularies, including those for which they do not have necessary licenses or permissions. It is also easy to exclude all source vocabularies that have particular restriction levels or all information in particular languages. In addition to identifying the source(s), restriction levels, and language of the information it contains, the Metathesaurus includes various more specific concept name flags and relationship labels that can help UMLS users to exclude content that is not relevant or helpful for particular applications.

MetamorphoSys, the install and customization program distributed with the UMLS (Section 6), makes it easy to generate custom subsets. MetamorphoSys also includes default settings that generate subsets that may be generally useful. MetamorphoSys can be also used to change the default preferred names of concepts (explained in Section 2.2.6); to change the default character set (from 7-bit ASCII to Unicode UTF8); and to include versioned vocabulary source abbreviations in every Metathesaurus file (see [section 2.1](#))

MetamorphoSys also generates special subsets referred to as Content Views. A content view may specify any pre-defined subset of the Metathesaurus that is useful for some specific purpose. The actual definition of a content view can take a variety of different forms: 1. An actual list of Metathesaurus UIs maintained over time. 2. A list of sources that participate in the view. 3. A complex query that identifies particular sets of data.

A Content View Flag (CVF) consists of an arbitrary bit field, with each bit representing membership in a particular content view; each content view is documented in MRDOC.RRF. The first Content View available in the 2005AA release, the MetaMap NLP View, identifies terms that are useful for Natural Language Processing. The CVF value in rows with these terms carry the value "1" in the "256" bit. MetamorphoSys users who wish to use this special subset should use the File Menu, Enable/Disable Views to implement this feature.

2.0.4 Metathesaurus Release Formats

Metathesaurus users may select from two relational formats: the Rich Release Format (RRF), introduced in 2004, and the Original Release Format (ORF). Both are available as output options of MetamorphoSys, the UMLS install and customization program (Section 6). All Rich Release Format file names have an extension (.RRF). Original Release Format files have no extension. Both formats are described in this documentation (usually abbreviated as RRF and ORF). There is also a White Paper explaining the rationale for the Rich Release Format and a detailed description of the differences between the .RRF files and the Original Format files.

The Rich Release Format has a number of advantages and is the preferred format for new users of the Metathesaurus and for most data creation applications.

2.1 Source Vocabularies

The Metathesaurus contains concepts, concept names, and other attributes from more than 100 terminologies, classifications, and thesauri, some in multiple editions. There is a concept in the Metathesaurus for each source vocabulary itself, which is assigned the semantic type "Intellectual Product". A special file (MRSAB.RRF and MRSAB in ORF) stores the version of each source vocabulary present in a particular edition of the Metathesaurus. All other Metathesaurus files that reference source vocabularies use "root" or versionless abbreviations, e.g., ICD9CM, not ICD9CM2003, thus avoiding routine wholesale updates to reflect the new versions. If you prefer to have versioned vocabulary source abbreviations in your custom Metathesaurus subset files, MetamorphoSys offers this as an option.

A complete list of the Metathesaurus source vocabularies with their root and versioned source abbreviations appears in Appendix B.4 of this documentation. The list is alphabetized by the abbreviation for that vocabulary source that is used in the Metathesaurus. Appendix B.4 includes other information including: the number of its concept names that are present in the Metathesaurus, the type of hierarchies or contexts it has (if any), and whether it is one of the small number of source

vocabularies that is not routinely updated in the Metathesaurus.

The Metathesaurus source vocabularies include terminologies designed for use in patient-record systems; large disease and procedure classifications used for statistical reporting and billing; more narrowly focused vocabularies used to record data related to psychiatry, nursing, medical devices, adverse drug reactions, etc.; disease and finding terminologies from expert diagnostic systems; and some thesauri used in information retrieval. A categorized list of the English-language source vocabularies is available.

2.1.1 Inclusion of U.S. Standard Code Sets and Terminologies

The Metathesaurus includes the code sets mandated for use in electronic administrative transactions in the U.S. under the provisions of the Health Insurance Portability and Accountability Act (HIPAA). With the exception of the National Drug Codes (NDC), the Metathesaurus includes all concepts and terms from these code sets. NDC codes available from the Food and Drug Administration are included as attributes of clinical drug concepts present in the FDA National Drug Code Directory (MTHFDA), which is a source vocabulary.

NLM intends to incorporate all clinical terminologies designated as target U.S. government-wide standards by the Consolidated Health Informatics (CHI) initiative and/or recommended as U.S. standards by the National Committee on Vital and Health Statistics. Several of these (e.g., LOINC, SNOMED CT, RxNorm) are already present in the Metathesaurus.

The fact that a vocabulary has been designated as a HIPAA or CHI standard is included in Appendix B.4.

2.1.2 Inclusion of Languages Other than English

The Metathesaurus structure can accommodate translations of its source vocabularies into languages other than English. Many translations in many different languages are present in this edition of the Metathesaurus. The Metathesaurus includes many translations of some source vocabularies, e.g., NLM's Medical Subject Headings (MeSH) and the International Classification of Primary Care; one or a few of others, and, in many cases, only the English version. As previously explained, MetamorphoSys (see Section 6) makes it easy to create a subset of the Metathesaurus that excludes the languages that are not relevant in a particular application.

2.2 Concepts, Concept Names, and their Identifiers

The Metathesaurus is organized by concept. One of its primary purposes is to connect different names for the same concept from many different vocabularies. The Metathesaurus assigns several types of unique, permanent identifiers to the concepts and concept names it contains, in addition to retaining all identifiers that are present in the source vocabularies. The Metathesaurus "concept structure" includes concept names, their identifiers, and key characteristics of these concept names (e.g., language, vocabulary source, name type). The entire concept structure appears in a single file in the Rich Release Format (MRCONSO.RRF). An abbreviated version of the concept structure is split between two files in the Original Format (MRCON and MRSO).

2.2.1 Concepts and Concept Identifiers

A concept is a meaning. A meaning can have many different names. A key goal of Metathesaurus construction is to understand the intended meaning of each name in each source vocabulary and to link all the names from all of the source vocabularies that mean the same thing (the synonyms). This is not an exact science. The construction of the Metathesaurus is based on the assumption that specially trained subject experts can determine synonymy with a degree of accuracy that is highly useful. Metathesaurus editors decide what view of synonymy to represent in the Metathesaurus concept structure. Please note that each source vocabulary's view of synonymy is also present in the Metathesaurus, irrespective of whether it agrees or disagrees with the Metathesaurus view.

Each concept or meaning in the Metathesaurus has a unique and permanent concept identifier (CUI). The CUI has no intrinsic meaning. In other words, you cannot infer anything about a concept just by looking at its CUI. In principle, the identifier for a concept never changes, irrespective of changes over time in the names that are attached to it in the Metathesaurus or in the source vocabularies.

In actuality, a CUI will be removed from the Metathesaurus when it is discovered that two CUIs actually name the same concept – in other words, when undiscovered synonymy comes to light. In these cases, one of the two CUIs will be retained, all relevant information in the Metathesaurus will be linked to it, and the other CUI will be retired.

Retired CUIs are never re-used. Each edition of the Metathesaurus includes files that detail any such changes from the previous edition. One Metathesaurus file (MRCUI.RRF and MRCUI in ORF) tracks such changes from 1991 to the present, allowing users to determine the fate of any CUI that is no longer present in the Metathesaurus.

2.2.2 Concept Names and String Identifiers

Each unique concept name or string in each language in the Metathesaurus has a unique and permanent string identifier (SUI). Any variation in character set, upper-lower case, or punctuation is a separate string, with a separate SUI. The same string in different languages (e.g., English and Spanish) will have a different string identifier for each language. If the same string, e.g., Cold, has more than one meaning, the string identifier will be linked to more than one concept identifier (CUI).

2.2.3 Atoms and Atom Identifiers

The basic building blocks or "atoms" from which the Metathesaurus is constructed are the concept names or strings from each of the source vocabularies. Each and every occurrence of a string in each source vocabulary is assigned a unique atom identifier (AUI). If exactly the same string appears twice in the same vocabulary, for example, as both the long name and the short name for the same concept or as an alternate name for two different concepts in the same vocabulary source, a unique AUI is assigned for each occurrence. When the same string appears in multiple source vocabularies, it will have AUIs for every time it appears as a concept name in each of those sources. All of these AUIs will be linked to a single string identifier (SUI), since they represent occurrences of the same string. Unlike string identifiers, a single AUI is always linked to a single concept identifier, because *each occurrence* of a string in a source can only have one meaning.

AUIs appear in the RRF (.RRF files), but not in the ORF.

2.2.4 "Terms" and Lexical Identifiers

For English language entries in the Metathesaurus only, each string is linked to all of its lexical variants or minor variations by means of a common term identifier (LUI). (In the Metathesaurus, therefore, an English "term" is the group of all strings that are lexical variants of each other.) English lexical variants are detected using the lvg program, one of the UMLS lexical tools (see Section 4). As similar tools become available for other languages, they may be used to create lexical variant groups in other languages. (In the meantime, the LUI for a non-English string is really another string identifier.)

Like a string identifier, the LUI for an English string may be linked to more than one concept. This occurs when strings that are lexical variants of each other have different meanings. In contrast, each string identifier and each atom identifier can only be linked to a single LUI.

2.2.5 Uses of Concept, String, Atom, and Term Identifiers

In the Metathesaurus, every CUI (concept) is linked to at least one AUI (atom), SUI (string), and LUI (term), but can be linked to many of each of these. Every AUI (atom) is linked to a single SUI (string), a single LUI (term), and a single CUI (concept). Each SUI (string) can be linked to many AUIs (atoms), to a single LUI (term), and to more than one CUI (concept) – although the typical case is one CUI. Each LUI (term) can be linked to many AUIs (atoms), many SUIs (strings), and more than one CUI (concept) – although the typical case is one CUI.

FIGURE 1.

Concept (CUI)	Terms (LUIs)	Strings (SUIs)	Atoms (AUIs) * RRF Only

C0004238	L0004238	S0016668	A0027665
Atrial Fibrillation (preferred) Atrial Fibrillations Auricular Fibrillation Auricular Fibrillations	Atrial Fibrillation (preferred) Atrial Fibrillations	Atrial Fibrillation (preferred)	Atrial Fibrillation (from MSH) A0027667 Atrial Fibrillation (from PSY)
		S0016669 Atrial Fibrillations	A0027668 Atrial Fibrillations (from MSH)
	L0004327 (synonym) Auricular Fibrillation Auricular Fibrillations	S0016899 Auricular Fibrillation (preferred)	A0027930 Auricular Fibrillation (from PSY)
		S0016900 (plural variant) Auricular Fibrillations	A0027932 Auricular Fibrillations (from MSH)

In the abbreviated example in Figure 1, "Atrial Fibrillation" appears as an atom in more than one source vocabulary and has a distinct AUI for each occurrence. Since each of these atoms has an identical string or concept name, they are linked to a single SUI. "Atrial Fibrillations", the plural of "Atrial Fibrillation" has a different string identifier. Since the singular and plural are lexical variants of each other, both are linked to the same LUI. There is a different LUI and different SUIs and AUIs for "Auricular Fibrillation" and its plural "Auricular Fibrillations." Since "Atrial Fibrillation" and "Auricular Fibrillation" have been judged to have the same meaning, they are linked to the same CUI.

All of these identifiers serve important purposes in building the Metathesaurus, in allowing efficient and accurate customization for specific purposes, and in identifying changes in its concept and concept name coverage over time.

CUIs link all information in the Metathesaurus related to particular concepts. In other words, a CUI can be used to retrieve all the concept names, relationships, and attributes for a particular concept that appear in any Metathesaurus file. CUIs also serve as permanent, publicly available identifiers for biomedical concepts or meanings to which many individual source vocabularies are linked. Users of the Metathesaurus are strongly encouraged to incorporate CUIs in their local applications – to support data exchange and linking and to assist migration between the use of individual source vocabularies should that become necessary in the future.

Users of the Metathesaurus are also encouraged to incorporate SUIs in local applications. Inclusion of SUIs will allow more efficient updating of local systems as new versions of the Metathesaurus are issued.

The value of retaining LUIs in local applications (as opposed to their use in creating the customized version of the Metathesaurus to be used locally) will vary depending on local system approaches to detecting and dealing with minor variations in language.

AUIs link all information in the Metathesaurus related to particular atoms or occurrences of strings in a specific source vocabulary. AUIs can assist users of the Metathesaurus in identifying those cases in which a source vocabulary's concept structure differs from that of the Metathesaurus. Many users of the Metathesaurus will have no need to store these identifiers in local applications.

2.2.6 Default Preferred Names for Metathesaurus Concepts

As a convenience for those who build the Metathesaurus, one string from one English term is designated and labeled as the *default* preferred name of each concept in the Metathesaurus.To avoid laborious selection among alternative terms and strings, selection of the default preferred name for any Metathesaurus concept is based on an order of precedence of all the types of English strings in all the Metathesaurus source vocabularies. Different types of strings, e.g., preferred terms, cross references, abbreviations, from each vocabulary will have different positions in this order. The factors considered in

establishing the default order of precedence include breadth of subject coverage, frequency of update, and the degree to which the source's concept names are used in regular clinical or biomedical discourse. The default order of precedence appears in MRRANK.RRF (MRRANK in ORF), and in Appendix B, Section B.5 of this documentation.

The default order of precedence will not be suitable for all applications of the Metathesaurus. MetamorphoSys (Section 6) can be used to change the selection of preferred names to feature terminology from the source vocabularies most appropriate to particular user populations. For example, concept names from SNOMED CT may be preferred in clinical applications, and terminology from MeSH may be preferred in literature retrieval systems.

2.2.7 Strings with Multiple Meanings

In some cases, the same name (with or without differences in upper-lower case) may apply to different concepts, usually (but not always) in different Metathesaurus source vocabularies. In the abbreviated example that follows, the string "Cold" is a name for the temperature in one vocabulary. In another vocabulary, "Cold" is an alternate name for the "Common cold". In a third vocabulary, "COLD" is an acronym for "chronic obstructive lung disease". As a result, "Cold" or "COLD" appears as a name of more than one concept in the Metathesaurus.

Representation of Ambiguity in the Metathesaurus:

Separate Metathesaurus files (AMBIGLUI.RRF and AMBIGSUI.RRF (AMBIG.LUI and AMBIG.SUI in ORF)) contain the LUIs and SUIs of all ambiguous terms and strings known to the Metathesaurus.

Prior to the 2005AA release, NLM created artificial strings, e.g., Cold <1>, to give each ambiguous meaning a unique name. These terms carried the source|term type value of MTH|MM, which has been removed from the Metathesaurus. The AM attributes used as explicit ambiguous string indicators were also removed.

FIGURE 2.

Concepts (CUIs)	Terms(LUIs)	Strings (SUIs)	Atoms (AUIs) ** RRF only
C0009264 cold temperature	L0215040 cold temperature	S0288775 cold temperature	A0318651 cold temperature (from CSP)
	L0009264 Cold Cold	S0007170 Cold	A0016032 Cold (from MTH)
		S0026353 Cold	A0040712 Cold (from MSH)
C0009443 Common Cold	L0009443 Common Cold	S0026747 Common Cold	A0041261 Common Cold (from MSH)
	L0009264 Cold Cold	S0007171 Cold	A0016033 Cold (from MTH)
		S0026353 Cold	A0040708 Cold (from COSTAR)

C0024117 Chronic Obstructive Airway Disease	L0498186 Chronic Obstructive Airway Disease	S0837575 Chronic Obstructive Airway Disease	A0896021 Chronic Obstructive Airway Disease (from MSH)
	L0008703 Chronic Obstructive Lung Disease	S0837576 Chronic Obstructive Lung Disease	A0896023 Chronic Obstructive Lung Disease (from MSH)
	L0009264 COLD COLD	S0829315 COLD	A0887858 COLD (from MTH)
		S0474508 COLD	A0539536 COLD (from SNMI)

2.2.8 Concept Names added during Metathesaurus Construction

Although the vast majority of concept names present in the Metathesaurus come from one or more of its source vocabularies, some concept names are created during Metathesaurus construction. This occurs in the following circumstances:

A unique name is created for a string with multiple meanings (the case explained in Section 2.2.7),

(a)   A more explicit name is created when none of the source vocabulary names for a concept conveys its meaning adequately,

(b)   An American English variant is generated for a British spelling,

(c)   An equivalent basic Latin ASCII character set string is generated for a string in an extended character set, such as Unicode.

Like all other concept names in the Metathesaurus, names created during Metathesaurus construction are labeled to indicate their source.

2.3 Relationships and Relationship Identifiers

The Metathesaurus includes many relationships between different concepts (in addition to the synonymous relationships in the Metathesaurus concept structure described in Section 2.2). Most of these relationships come from individual source vocabularies. Some are added by NLM during Metathesaurus construction. Some have been contributed by Metathesaurus users to support certain types of applications.

Relationships are expressed in terms of CUIs (in the RRF and ORF) and AUIs (in the RRF only). Metathesaurus relationship files do not include concept names.

In general, the Metathesaurus indicates the author of each relationship, that is, one of the source vocabularies, the Metathesaurus itself, or another supplier. Some relationships added in the early years of Metathesaurus development (less than 6% of the current total and declining) are attributed to the Metathesaurus, but actually came from specific source vocabularies.

2.3.1 Basic categories of non-synonymous relationships

The Metathesaurus contains non-synonymous relationships between concepts from the same source vocabulary (*intra*-source vocabulary relationships) and between concepts in different vocabularies (*inter*-source vocabulary relationships). **The Metathesaurus does *not* include *all possible* non-synonymous relationships between the concepts it contains.** It includes all relationships present in its source vocabularies and some additional relationships designed to connect related concepts. In general, the relationships asserted by source vocabularies connect closely related concepts, such as those that share some common property or are related by definition. For example, a member of a class of drugs (e.g., penicillin) will be connected to the name for the class (e.g., antibiotics); a bacterial infection will be connected to the bacterium that causes it.

2.3.1.1 *Intra*-Source Relationships

The majority of *intra*-source relationships are asserted or implied by the individual source vocabularies. Such relationships occur in a source vocabulary's explicit or implied hierarchical arrangements or contexts, cross-reference structures, rules for applying qualifiers, or connections between different types of names for the same concept (e.g., abbreviations and full forms). The primary Metathesaurus relationships file, that is, MRREL.RRF and MRREL in the ORF contains the "distance -1" hierarchical relationships, i.e., immediate parents, immediate child, and immediate sibling relationships, as well as other types of intra-source relationships.

A subset of the contextual or hierarchical relationships are also distributed in a special contexts file (MRCXT.RRF and MRCXT in ORF) to facilitate the construction of user displays. A "computable" representation of the complete hierarchies is provided in MRHIER.RRF only. MRHIER.RRF, for example, represents all sibling relationships even when there are thousands of siblings. Appendix B.4 indicates which source vocabularies have hierarchical contexts, which of these allow concepts to appear in multiple hierarchies, and whether sibling relationships are represented in MRCXT.RRF and MRCXT in ORF or only in MRHIER.RRF.

ORF users may omit MRCXT if they don't want these selected, pre-computed contexts.

Some of the intra-source vocabulary relationships are statistical relationships, which are computed by determining the frequency with which concepts in specific vocabularies co-occur in records in a database. For example, there are co-occurrence relationships for the number of times concepts have co-occurred as key topics within the same articles, as evidenced by the Medical Subject Headings assigned to those articles in the MEDLINE database. Co-occurrence relationships have been also computed for different ICD-9-CM diagnosis codes assigned to the same patients as reflected in a discharge summary database. In contrast to the relationships asserted within source vocabularies, the statistical relationships in the Metathesaurus can connect very different concepts, such as diseases and drugs. There are specific Metathesaurus files for the co-occurrence relationships (MRCOC.RRF and MRCOC in ORF).

2.3.1.2 *Inter*-Source Relationships

The primary *inter*-source relationships in the Metathesaurus are the synonymous relationships represented in the Metathesaurus concept structure (Section 2.2). The Metathesaurus also includes some relationships between non-synonymous concepts from different source vocabularies. Some of these inter-source relationships are generated during Metathesaurus construction to connect specific "orphan" concepts (with few or no ancestors, siblings, or children in their own source vocabularies) to the richer contextual information in another source vocabulary. Some are supplied by Metathesaurus users who find "like" or "similar" relationships a useful addition to the Metathesaurus's relatively strict view of synonymy. In both cases, these relationships are distributed in MRREL.RRF and MRREL in ORF.

Many inter-source relationships between non-synonymous concepts are produced through specific efforts to create a mapping between two different source vocabularies. These mappings may be created by an individual source vocabulary producer, by a third party with a particular need for a mapping, or by NLM or under NLM supervision specifically for distribution within the Metathesaurus. The number of NLM-supervised mappings is expected to increase. There are specific Metathesaurus files for mappings in the RRF (MRMAP.RRF and MRSMAP.RRF). A subset of the mappings appear in MRATX in the ORF. Mappings involving SNOMED CT appear in the RRF only.

2.3.2 Relationship Labels

All relationships (outside the basic concept structure) in the Metathesaurus carry a general label (REL), describing their basic

nature, such as Broader, Narrower, Child of, Qualifier of, etc and are identified by their source. Most of these relationships are either directly asserted in a source vocabulary or are implied by the structure of the source vocabulary. A complete list of the general relationship labels appears as MRDOC.RRF in Appendix B.3 in this documentation.

About a quarter of the relationships in the Metathesaurus *also* carry an additional label (RELA), obtained from a source vocabulary, that explains the nature of the relationship more exactly, such as is_a, branch_of, component_of. The Digital Anatomist vocabulary and RxNorm are examples of source vocabularies that include such relationship labels. A complete list of the additional relationship labels appears in MRDOC.RRF and in Appendix B.3 in this documentation.

2.3.3 Relationship Identifiers

Every relationship present in the Metathesaurus has a unique relationship identifier (RUI). The primary purpose of these identifiers is to enable easy detection of changes in relationships across versions of the Metathesaurus. The appearance or disappearance of a relationship identifier indicates a change in the relationships present in the Metathesaurus.

Some source vocabularies have their own relationship identifiers. Where they exist, these identifiers are also present in the Metathesaurus.

2.4 Attributes and Attribute Identifiers

In the Metathesaurus, attributes include every discrete piece of information about a concept, an atom, or a relationship that is not (1) part of the basic Metathesaurus concept structure (Section 2.2) or (2) distributed in one of the relationship files (Section 2.3).

2.4.1 Kinds of Attributes

The Metathesaurus includes concept attributes, atom attributes, and relationship attributes.

Concept attributes are added during Metathesaurus construction and apply to all names of a concept. For example, the semantic types "Pathologic Function" and "Finding" are attributes of the concept with the preferred name "Atrial Fibrillation" and are applicable to any atom connected to that concept.

Atom attributes come from a particular source vocabulary. Some of them are of general interest; others are relevant only to a particular source vocabulary. For example, the definition "Disorder of cardiac rhythm characterized by rapid, irregular atrial impulses and ineffective atrial contractions." is an attribute of the atom "Atrial Fibrillation" that comes from the Medical Subject Headings (MeSH). It may be one of several definitions connected to names of this concept, because the Metathesaurus includes all definitions provided by any of its source vocabularies. Although this particular definition comes from MeSH, it might well be useful in Metathesaurus applications that otherwise do not use MeSH. In contrast, the date an occurrence of a string (an atom) was added to a source vocabulary applies only to that specific atom. The utility of specific atom attributes will vary considerably for different applications of the Metathesaurus.

Relationship attributes come from a particular source vocabulary and describe special characteristics of particular relationships in that source, e.g., refinability.

The majority of attributes are distributed in MRSAT.RRF and MRSAT in the ORF. In these files, each row contains the name of the attribute, the source of the attribute, and the value of the attribute, in addition to all appropriate identifiers. There are separate files for selected attributes such as the semantic types (MRSTY.RRF and MRSTY in the ORF) and the definitions (MRDEF.RRF and MRDEF in the ORF).

2.4.2 Attribute Identifiers

Each occurrence of each attribute within the Metathesaurus is assigned a unique attribute identifier (ATUI). The appearance or disappearance of ATUIs signals changes in the content of the Metathesaurus, thus ATUIs assist the efficient production of a complete change set for each new version of the Metathesaurus. ATUIs appear only in the RRF, not in the ORF.

2.5 Data About the Metathesaurus

The Metathesaurus contains a number of files that provide useful "metadata" or data about the Metathesaurus itself. The metadata files describe (1) characteristics of the current version of the Metathesaurus; (2) changes between the current version and the previous version; and (3) the history of concept identifiers (CUIs) from 1991 to the present.

2.5.1 Characteristics of the Current Metathesaurus

There are discrete Metathesaurus files for:

a)   the names and sizes of every Metathesaurus file (MRFILES.RRF and MRFILES in ORF),

b)   the names and size range of every Metathesaurus data element (MRCOLS.RRF and MRCOLS in ORF),

c)   the possible values for selected data elements that contain a finite set of abbreviated values (MRDOC.RRF only). NOTE: eventually this file will include values for <u>every</u> data element that contains a finite set of abbreviated values

d)   the source vocabularies in the Metathesaurus (MRSAB.RRF and MRSAB in ORF),

e)   the LUIs and SUIs for terms and strings that are known to be ambiguous, that is, to have multiple meanings (to be linked to multiple concept identifiers) within the Metathesaurus (AMBIGLUI.RRF and AMBIGSUI.RRF in RRF and AMBIGLUI and AMBIGSUI in ORF),

f)   the order of precedence of vocabulary source and term types that is used to compute the default preferred concept name for each concept in the Metathesaurus (MRRANK.RRF and MRRANK in ORF). NOTE: MetamorphoSys can be used to change this order.

MRCOLS, MRDOC, MRSAB, and MRRANK contain data that do not appear in the actual Metathesaurus content files. The others are computable from the Metathesaurus content files. They are pre-computed and provided in separate files as a convenience to users of the Metathesaurus.

2.5.2 Changes between the Current Metathesaurus and the Previous Version

Each version of the Metathesaurus contains a set of files that summarize changes from the previous version.

CHANGE/MERGEDCUI.RRF in the RRF (CHANGE/MERGED.CUI in the ORF) documents cases in which two discrete concepts in the previous version of the Metathesaurus are now considered to be synonyms.

CHANGE/MERGEDLUI.RRF in the RRF (CHANGE/MERGED.LUI in the ORF) documents cases in which two discrete terms in the previous version of the Metathesaurus are now identified as lexical variants of each other, based on the current version of luinorm (the program used to compute them).

Three files contain the CUIs, LUIs, and SUIs for Metathesaurus concepts, terms, and strings that appeared in the previous version, but are not in the current version (CHANGE/DELETEDCUI.RRF, CHANGE/DELETEDLUI.RRF, CHANGE/ DELETEDSUI.RRF in the RRF and CHANGE/DELETED.CUI, CHANGE/DELETED.LUI, CHANGE/DELETED.SUI in the ORF).

NOTE: In future versions of the Metathesaurus change files will also be provided for relationships and attributes in the RRF only. The generation of these files is dependent on the relationship and attribute identifiers (RUI and ATUI) introduced in the 2004AA version of the Metathesaurus.

2.5.3 Historical CUIs

The retired CUI file (MRCUI.RRF in RRF and MRCUI in ORF) includes all CUIs present in any previous version of the

Metathesaurus, but not in the current version. In general, the file maps the retired CUI to one or more current CUIs.

2.6 Concept Name Indexes

2.6.0 Introduction

To assist system developers in building applications that retrieve all strings or concept names which include specific words or groups of words, three indexes to the concept names are provided: a Word Index, a Normalized Word Index (for English words only), and a Normalized String Index (for English strings only). The indexes are described in sections 2.6.1, and 2.6.3 respectively. To make the distinctions among them clearer, the examples include words or strings that would appear in each index for the following set of Metathesaurus concept names:

Lung Diseases, Obstructive	(C0024117, L0024117, S0058463)
Obstructive Lung Diseases	(C0024117, L0024117, S0068169)
Lung Disease, Obstructive	(C0024117, L0024117, S0058458)
Obstructive Lung Disease	(C0024117, L0024117, S0068168)

2.6.1 Word Index

2.6.1.1 Description

The word index connects each individual word in any Metathesaurus string to all its related string, term, and concept identifiers. There are separate word index files for each language in the Metathesaurus.

There is one entry for each word found in each unique string in each language. Each entry has five subelements.

1. LAT - 3-letter abbreviation for language
2. WD - Word
3. CUI - concept unique identifier
4. LUI - term identifier
5. SUI - string identifier

Sample records:

ENG|000003|C1273274|L3139159|S3660797|
ENG|000003|C1306276|L3139160|S3660798|

2.6.1.2 Definition of a Word

For the purpose of creating this index, a word is defined as a token containing only alphanumeric characters with length one or greater; for more information, see the SPECIALIST Lexicon and tools.

2.6.1.3 Word Index Example

For the four concept names listed in Section 2.6.0, the word index will contain multiple entries for each of the following words: disease, diseases, lung, obstructive. Two of the entries generated for the names "Lung Disease, Obstructive" and "Obstructive Lung Disease" are shown below:

ENG|disease|C0024117|L0024117|S0058458|
ENG|disease|C0024117|L0024117|S0068168|

2.6.2 Normalized Word Index

2.6.2.1 Description

The normalized word index connects each individual normalized English word to all its related string, term, and concept identifiers.

There is one entry for each normalized word found in each unique English string. There are no entries for other languages in this index. Each entry has five subelements.

1. LAT - (always ENG in this edition of the Metathesaurus)
2. NWD - normalized word
3. CUI - concept unique identifier
4. LUI - term unique identifier
5. SUI - string identifier

2.6.2.2 Definition of Normalized Word

The normalization process involves breaking a string into its constituent words, lowercasing each word and converting it to its uninflected form. Normalized words are generated by uninflecting each word and stripping out a small number of stop words. The uninflected forms are generated using the SPECIALIST lexicon if the words appear in the lexicon; otherwise they are generated algorithmically.

2.6.2.3 Normalized Word Example

For the four concept names listed in Section 2.6.0 the normalized word index will contain multiple entries for each of the following words: disease, lung, obstructive. Since the normalized word index contains base forms only, it does not contain entries for the plural "diseases". In this index, therefore, all four concept names are linked to the normalized word "disease", as follows:

ENG|disease|C0024117|L0024117|S0058458|
ENG|disease|C0024117|L0024117|S0058463|
ENG|disease|C0024117|L0024117|S0068168|
ENG|disease|C0024117|L0024117|S0068169|

2.6.3 NORMALIZED STRING INDEX

2.6.3.1 Description

The normalized string index connects the normalized form of a Metathesaurus string to all its related string, term, and concept identifiers. There is one entry for each unique (non-normalized) English string. There are no entries for other languages in this index. Each entry has five subelements.

1. LAT (always ENG in this edition of the Metathesaurus)
2. NSTR - normalized string
3. CUI - concept unique identifier
4. LUI - term identifier
5. SUI - string identifier

2.6.3.2 Definition of Normalized String

The normalization process involves breaking a string into its constituent words, lowercasing each word, converting each word to its uninflected form, and sorting the words in alphabetic order. Normalized strings are generated by uninflecting each word leaving out a small number of stop words. The uninflected forms are generated using the SPECIALIST lexicon if the words

appear in the lexicon; otherwise they are generated algorithmically.

2.6.3.3 Normalized String Example

Since the four concept names listed in Section 2.6.0 are composed of the same set of normalized words, the Normalized String Index will contain four entries for a single string: disease lung obstructive, in which the component normalized words appear in alphabetical order. The **complete** set of Normalized String Index entries generated by the four concept names is as follows:

ENG|disease lung obstructive|C0024117|L0024117|S0058458|
ENG|disease lung obstructive|C0024117|L0024117|S0058463|
ENG|disease lung obstructive|C0024117|L0024115|S0068168|
ENG|disease lung obstructive|C0024117|L0024117|S0068169|

2.6.4 Word Index Programs

The programs that generate these indexes are written in Java. They may be of use to system developers who are developing their own interfaces to the UMLS data or for other purposes. Section 4 includes information about these and other lexical programs provided with the UMLS Knowledge Sources.

2.7 File Formats - Metathesaurus Rich Release Format (RRF) and Original Release Format (ORF)

2.7.0 Introduction

Metathesaurus users may select from two relational formats: the Rich Release Format (RRF), first introduced in 2004, and the Original Release Format (ORF). Both are available as output options of MetamorphoSys, the UMLS install and customization program (Section 6).

Developers are encouraged to use the RRF, which offers significant advantages in source vocabulary "transparency" (that is, ability to represent the detailed semantics of each source vocabulary exactly); in the ability to generate complete and accurate change sets between versions of the Metathesaurus; and in more convenient representations of concept name, source, and hierarchical context information. A more complete discussion of the rationale for the RRF and a detailed description of the differences between the two formats are available.

Neither Metathesaurus format is fully normalized. By design, there is duplication of data among different files and within certain files. In particular, relationships between different Metathesaurus concepts appear twice (e.g., from entry A to entry B and from entry B to entry A). Developers will need to make their own decisions about the extent to which this redundancy should be retained, reduced, or increased for their specfic applications.

Section 2.7.1 describes the files in the RRF. Section 2.7.2 describes the files in the ORF.

2.7.1 Metathesaurus Rich Release Format (RRF)

All file names begin with the letters MR (Metathesaurus Relational) and are followed by letters that denote the file contents (e.g., MRREL=relationships, MRSAB=source abbreviations), and then a file extension .RRF.

All files except MRRANK.RRF are sorted by row.

2.7.1.1 Data Files

The data in each Metathesaurus entry may be represented in more than 20 different "relations" or files. These files correspond to the four logical groups of data elements described in Sections 2.2- 2.5 and the indexes described in Section 2.6 as follows:

Concepts, Concept Names, and their sources (2.2) = MRCONSO.RRF

Attributes (2.3) = MRSAT.RRF, MRDEF.RRF, MRSTY.RRF, MRHIST.RRF

Relationships (2.4) = MRREL.RRF, MRCOC.RRF, MRCXT.RRF, MRHIER.RRF, MRMAP.RRF, MRSMAP.RRF

Data about the Metathesaurus (2.5) = MRFILES.RRF, MRCOLS.RRF, MRDOC.RRF, MRRANK.RRF, MRSAB.RRF, AMBIGLUI.RRF, AMBIGSUI.RRF, CHANGE/MERGEDCUI.RRF, CHANGE/MERGEDLUI.RRF, CHANGE/DELETEDCUI.RRF, CHANGE/DELETEDLUI.RRF, CHANGE/DELETEDSUI.RRF, MRCUI.RRF

Indexes (2.6) = MRXW_BAQ.RRF, MRXW_DAN.RRF, MRXW_DUT.RRF, MRXW_ENG.RRF, MRXW_FIN.RRF, MRXW_FRE.RRF, MRXW_GER.RRF, MRXW_HEB.RRF, MRXW_HUN.RRF, MRXW_ITA.RRF, MRXW_NOR.RRF, MRXW_POR.RRF, MRXW_RUS.RRF, MRXW_SPA.RRF, MRXW_SWE.RRF, MRXNW_ENG.RRF, MRXNS_ENG.RRF

2.7.1.2 Columns and Rows

Each file or named table of data values has by definition a fixed number of columns; the number of rows depends on the content of a particular version of the Metathesaurus.

A column is a sequence of all the values in a given data element or logical subelement. In general, columns for longer variable length data elements will appear to the right of columns for shorter and/or fixed length data elements. The information for all columns in the files is described in MRCOLS.RRF and in Appendix B.1.1, Metathesaurus Column Descriptions.

A row contains the values for one or more data elements or logical subelements for one Metathesaurus entry. Depending on the nature of the data elements involved, each Metathesaurus entry may have one or more rows in a given file. The values for the different data elements or logical subelements represented in the row are separated by vertical bars (|). If an optional element is blank, the vertical bars are still used to maintain the correct positioning of the subsequent elements. Each row is terminated by a vertical bar and line termination.

2.7.1.3 Descriptions of Each File

The descriptions of the files appear in the following order:
a) Key data about the Metathesaurus: Files; Columns or data elements; Documentation that explains the meaning of abbreviations that appear as values in Metathesaurus data elements and attributes,
b) Concept names and their vocabulary sources
c) Attributes
d) Relationships
e) Other data about the Metathesaurus
f) Indexes

Each file description lists the columns or data elements that appear in the file and includes sample rows from the file.

2.7.1.3.1 Files (File = MRFILES.RRF)

There is exactly one row in this file for each physical segment of each logical file. Data elements that appear in multiple files, e.g., CUI, AUI, will have multiple rows in this file.

Col.	Description
FIL	Physical FILENAME
DES	Descriptive Name
FMT	Comma separated list of column names (COL), in order
CLS	# of COLUMNS

RWS	# of ROWS
BTS	Size in bytes in this format (ISO/PC or Unix)

Sample Records

MRCOC.RRF|Co-occurringConcepts|CUI1,AUI1,CUI2,AUI2,SAB,COT,COF,COA,CVF|9|13939548|786509996|
MRSTY.RRF|Semantic Types|CUI,TUI,STN,STY,ATUI,CVF|6|1146352|64528811|

2.7.1.3.2 Data Elements (File = MRCOLS.RRF)

There is exactly one row in this file for each column or data element in each file. Data elements that appear in multiple files, e. g., CUI, AUI, will have multiple rows in this file.

Col.	Description
COL	Column or data element name
DES	Descriptive Name
REF	Documentation Section Number
MIN	Minimum Length, Characters
AV	Average Length
MAX	Maximum Length, Characters
FIL	Physical FILENAME in which this field occurs
DTY	SQL-92 data type for this column

Sample Records

AUI|Unique identifier for atom||8|8.00|8|MRCONSO.RRF|char(8)|
CODE|Unique Identifier or code for string in source||1|6.4|21|MRCONSO.RRF|varchar(50)|

2.7.1.3.3 Documentation for Abbreviated Values (File = MRDOC.RRF)

There is exactly one row in this table for each allowed value of selected data elements or attributes that have a finite number of abbreviations as allowed values. Examples of such data elements include TTY, ATN, TS, STT, REL, RELA.

Col.	Description
KEY	Data element or attribute
VALUE	Abbreviation that is one of its values
TYPE	Type of information in EXPL column
EXPL	Explanation of VALUE

Sample Records

ATN|DDF|expanded_form|Drug Doseform|
ATN|DHJC|expanded_form|HCPCS J-code|

2.7.1.3.4 Concept Names and Sources (File = MRCONSO.RRF)

There is exactly one row in this file for each atom (each occurrence of each unique string or concept name within each source vocabulary) in the Metathesaurus, i.e., there is exactly one row for each unique AUI in the Metathesaurus. Every string or concept name in the Metathesaurus appears in this file, connected to its language, source vocabularies, and its concept identifier. The values of TS, STT, and ISPREF reflect the default order of precedence of vocabulary sources and term types in MRRANK.RRF.

Col.	Description
CUI	Unique identifier for concept
LAT	Language of Term
TS	Term status
LUI	Unique identifier for term
STT	String type
SUI	Unique identifier for string
ISPREF	Atom status - preferred (Y) or not (N) for this string within this concept
AUI	Unique identifier for atom
SAUI	Source asserted atom identifier [optional]
SCUI	Source asserted concept identifier [optional]
SDUI	Source asserted descriptor identifier [optional]
SAB	Source abbreviation
TTY	Term type in source
CODE	"Most useful" source asserted identifier (if the source vocabulary has more than one identifier), or a Metathesaurus-generated source entry identifier (if the source vocabulary has none.)
STR	String
SRL	Source Restriction Level
SUPPRESS	Suppressible flag. Values = O, E, Y, or N.
	O: all obsolete content, whether they are obsolesced by the source or by NLM. These will include all atoms having obsolete TTYs, and other atoms becoming obsolete that have not acquired an obsolete TTY (eg. RxNorm SCDs no longer associated with current drugs, LNC atoms derived from obsolete LNC concepts) E: non-obsolete content marked suppressible by an editor. These do not have a suppressible SAB/ TTY combination. Y: non-obsolete content deemed suppressible during inversion. These can be determined by a specific SAB/TTY combination explicitly listed in MRRANK N: none of the above
	Default suppressibility as determined by NLM (i.e., no changes at the Suppressibility tab in MetamorphoSys) should be used by most users, but may not be suitable in some specialized applications. See the MetamorphoSys documentation (Section 6) for information on how to change the SAB/TTY suppressibility to suit your requirements. NLM strongly recommends that users not alter editor-assigned suppressibility, and MetamorphoSys cannot be used for this purpose.
CVF	Content view flag Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use.

Sample Records

C0001175|ENG|P|L0001175|VO|S0010340|Y|A0019182||M0000245|D000163|MSH|PM|D000163|
    Acquired Immunodeficiency Syndromes|0|N||
C0001175|ENG|S|L0001842|PF|S0011877|N|A2878223|103840012|62479008||SNOMEDCT|PT|62479008|AIDS|4|N||
C0001175|ENG|P|L0001175|VC|S0354232|Y|A2922342|103845019|62479008||SNOMEDCT|SY|62479008|
    Acquired immunodeficiency syndrome|4|Y||
C0001175|FRE|P|L0162173|PF|S0226654|Y|A0248753||||INS|MH|d000163|SIDA|3|N||
C0001175|RUS|P|L0904943|PF|S1108760|Y|A1165232||||RUS|MH|D000163|SPID|3|N||

2.7.1.3.5 Simple Concept and Atom Attributes (File = MRSAT.RRF)

There is exactly one row in this table for each concept, atom, or relationship attribute that does not have a sub-element structure. All Metathesaurus concepts and a minority of Metathesaurus relationships have entries in this file. This file includes all source vocabulary attributes that do not fit into other categories.

Col.	Description
CUI	Unique identifier for concept (if METAUI is a relationship identifier, this will be CUI1 for that relationship)
LUI	Unique identifier for term (optional - present for atom attributes, but not for relationship attributes)
SUI	Unique identifier for string (optional - present for atom attributes, but not for relationship attributes)
METAUI	Metathesaurus atom identifier (will have a leading A) or Metathesaurus relationship identifier (will have a leading R) or blank if it is a concept attribute.
STYPE	The name of the column in MRCONSO.RRF or MRREL.RRF that contains the identifier to which the attribute is attached, e.g., SAUI, SCUI, SRUI, CODE, CUI, AUI. Many attributes currently shown as linked to Metathesaurus AUIs will be linked to one of the source vocabulary identifiers as vocabularies that were added to the Metathesaurus prior to the development of the RRF are updated and brought into complete alignment with the RRF.
CODE	"Most useful" source asserted identifier (if the source vocabulary contains more than one) or a Metathesaurus-generated source entry identifier (if the source vocabulary has none). Optional -- present if METAUI is an AUI.
ATUI	Unique identifier for attribute
SATUI	Source asserted attribute identifier (optional - present if it exists)
ATN	Attribute name. Possible values appear in MRDOC.RRF and are described in Appendix B.2
SAB	Abbreviation of the source of the attribute. Possible values appear in MRSAB.RRF and are listed in Appendix B.4
ATV	Attribute value described under specific attribute name in Appendix B.2. A few attribute values exceed 1,000 characters. Many of the abbreviations used in attribute values are explained in MRDOC.RRF and included in Appendix B.3.
SUPPRESS	Suppressible flag. Values = O, E, Y, or N. Reflects the suppressible status of the attribute; not yet in use. See also SUPPRESS in MRCONSO.RRF and MRDEF.RRFand MRREL.RRF.
CVF	Content view flag Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use.

Sample Records

C0001175|L0001175|S0010339|A0019180|AUI|D000163|AT15797077||FX|MSH|AIDS Dementia Complex|N||

C0001175|L0001175|S0354232|A2922342|SAUI|62479008|AT34794876||DESCRIPTIONSTATUS|SNOMEDCT|0|N||
C0001175|L2810384|S3645548|A3814219|SCUI|62479008|AT33494582||CTV3ID|SNOMEDCT|XE0RX|N||
C0001175|L2810384|S3645548|A3814219|SCUI|62479008|AT33652930|\ISPRIMITIVE|SNOMEDCT|0|N||
C0001175|||R19334287|SRUI||AT37098279||REFINABILITY|SNOMEDCT|1|N||

2.7.1.3.6 Definitions (File = MRDEF.RRF)

There is exactly one row in this file for each definition in the Metathesaurus. A definition is an attribute of an atom (an occurrence of a string in a source vocabulary). A few approach 3,000 characters in length.

Col.	Description
CUI	Unique identifier for concept
AUI	Unique identifier for atom
ATUI	Unique identifier for attribute
SATUI	Source asserted attribute identifier [optional-present if it exists]
SAB	Abbreviation of the source of the definition
DEF	Definition
SUPPRESS	Suppressible flag. Values = O, E, Y, or N. Reflects the suppressible status of the attribute; not yet in use. See also SUPPRESS in MRCONSO.RRF and MRDEF.RRFand MRREL.RRF.
CVF	Content view flag Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use.

Sample Records

C0001175|A0019180|AT15060425||MSH|An acquired defect of cellular immunity associated with infection by the human immunodeficiency virus (HIV), a CD4-positive T-lymphocyte count under 200 cells/microliter or less than 14% of total lymphocytes, and increased susceptibility to opportunistic infections and malignant neoplasms. Clinical manifestations also include emaciation (wasting) and dementia. These elements reflect criteria for AIDS as defined by the CDC in 1993.|N||

C0001175|A0021048|AT14042185||CSP|one or more indicator diseases, depending on laboratory evidence of HIV infection (CDC); late phase of HIV infection characterized by marked suppression of immune function resulting in opportunistic infections, neoplasms, and other systemic symptoms (NIAID).|N||

C0001175|A0021055|AT18420297||PDQ|Acquired immunodeficiency syndrome. An acquired defect in immune system function caused by human immunodeficiency virus 1 (HIV-1). AIDS is associated with increased susceptibility to certain cancers and to opportunistic infections, which are infections that occur rarely except in individuals with weak immune systems.|N||

2.7.1.3.7 Semantic Types (File = MRSTY.RRF)

There is exactly one row in this file for each Semantic Type assigned to each concept. All Metathesaurus concepts have at least one entry in this file. Many have more than one entry. The TUI, STN, and STY are all direct links to the UMLS Semantic Network (Section 3).

Col.	Description
CUI	Unique identifier of concept
TUI	Unique identifier of Semantic Type
STN	Semantic Type tree number
STY	Semantic Type. The valid values are defined in the Semantic Network.
ATUI	Unique identifier for attribute
CVF	Content view flag
	Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use.

Sample Record

C0001175|T047|B2.2.1.2.1|Disease or Syndrome|AT17683839||

2.7.1.3.8.a Locators (File = MRLO.RRF)

This file has been deleted from Metathesaurus effective with the 2004AB release. Some of the information was outdated, some was duplicative of information contained in other Metathesaurus files, and some was easily obtained from other publicly available sources, e.g., PubMed.

2.7.1.3.8.b History (File = MRHIST.RRF)

This file tracks source-asserted history information. It currently includes SNOMED CT history only.

Col.	Description
CUI	Unique identifier for concept
SOURCEUI	Source asserted unique identifier
SAB	Source abbreviation
SVER	Release date or version number of a source
CHANGETYPE	Source asserted code for type of change
CHANGEKEY	CONCEPSTATUS (if history relates to a SNOMED CT concept) or DESCRIPTIONSTATUS (if history relates to a SNOMED CT atom)
CHANGEVAL	CONCEPSTATUS value or DESCRIPTIONSTATUS value after the change took place [NOTE: the change may have affected something other than the status value)
REASON	Explanation of change if present
CVF	Content view flag
	Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use.

Sample Records

C0000294|108821000|SNOMEDCT|20001101|0|CONCEPTSTATUS|0|||
C0000294|108821000|SNOMEDCT|20020731|2|CONCEPTSTATUS|0|FULLYSPECIFIEDNAME CHANGE||
C0000294|1185494016|SNOMEDCT|20020731|0|DESCRIPTIONSTATUS|0|||

C0000294|1461100014|SNOMEDCT|20030131|0|DESCRIPTIONSTATUS|0|||

2.7.1.3.9 Related Concepts (File = MRREL.RRF)

There is one row in this table for each relationship between concepts or atoms known to the Metathesaurus, with the following exceptions found in other files: co-occurrences found in MRCOC.RRF, and pair-wise mapping relationships between two source vocabularies found in MRMAP.RRF and MRSMAP.RRF.

Note that for asymmetrical relationships there is one row for each direction of the relationship. Note also the direction of REL - the relationship which the SECOND concept or atom (with Concept Unique Identifier CUI2 and Atom Unique Identifier AUI2) HAS TO the FIRST concept or atom (with Concept Unique Identifier CUI1 and Atom Unique Identifier AUI1).

Col.	Description
CUI1	Unique identifier of first concept
AUI1	Unique identifier for first atom
STYPE1	The name of the column in MRCONSO.RRF that contains the identifier used for the first concept or first atom in source of the relationship.
REL	Relationship of second concept or atom to first concept or atom
CUI2	Unique identifier of second concept
AUI2	Unique identifier for second atom
STYPE2	The name of the column in MRCONSO.RRF that contains the identifier used for the second concept or second atom in the source of the relationship.
RELA	Additional (more specific) relationship label (optional)
RUI	Unique identifier for relationship
SRUI	Source asserted relationship identifier, if present
SAB	Abbreviation of the source of relationship
SL	Source of relationship labels
DIR	Source asserted directionality flag. Y indicates that this is the direction of the relationship in its source; N indicates that it is not; a blank indicates that it is not important or has not yet been determined.
SUPPRESS	Suppressible flag. Values = O, Y, E, or N. Reflects the suppressible status of the relationship; not yet in use. See also SUPPRESS in MRCONSO.RRF and MRDEF.RRF and MRREL.RRF.
CVF	Content view flag Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use.

Sample Records

C0002372|A0022284|AUI|RB|C0002371|A0022279|AUI||R01983351||MSH|MSH|||N||
C0002372|A0022284|AUI|SY|C0002372|A0062352|AUI||R18851331||MSH|MSH|||N||

2.7.1.3.10 Co-occurring Concepts (File = MRCOC.RRF)

This file includes statistical aggregations of co-occurrences of meanings in external data sources. These exist at the AUI level. There are two rows in this table for each pair of atoms that co-occur in each information source represented: one for each direction of the relationship. (Note that the COA data may be different for each direction of the relationship.) Many Metathesaurus concepts have no entries in this file. Due to the very large number of co-occurrence relationships, they are

distributed in a separate file.

Col.	Description
CUI1	Unique identifier of first concept
AUI1	Unique identifier of first atom
CUI2	Unique identifier of second concept or not present Note: Where CUI2 is not present and COT is LQ (MeSh topical qualifier), the count of citations of CUI1 with no MeSH qualifiers is reported in COF.
AUI2	Unique identifier of second atom
SAB	Abbreviation of the Source of co-occurrence information
COT	Type of co-occurrence
COF	Frequency of co-occurrence, if applicable
COA	Attributes of co-occurrence, if applicable
CVF	Content view flag Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use.

Co-occurrences are concepts that occur together in the same "entries" in some information source. The relationships represented here are obtained from machine-manipulation of the information source. Co-occurrence relationships may exist between similar concepts (e.g., "Atrial Fibrillation" and "Arrhythmia") or between very different concepts that nevertheless have some important connection in the field of biomedicine (e.g., "Atrial Fibrillation" and "Digoxin"), or between a primary concept and a qualifier e.g., "Lithotripsy" and "instrumentation". A co-occurrence relationship can exist between two concepts that have no other apparent relationship, although the frequency of such co-occurrences will be small.

In the current Metathesaurus, there are three sources of co-occurrence data: MEDLINE, AI/RHEUM, and CCPSS. From MEDLINE, co-occurrence data was computed for concepts that were designated as principal or main points in the same journal article i.e., the co-occurrence counts do not include articles in which either or both of the concepts were present and indexed in MEDLINE but not designated as main points. (A concept is considered to be a main point if the * is attached to the main heading or any of its subheadings.)

Two overall frequencies of MEDLINE co-occurrence are provided: one for recent MEDLINE data (MED) and one for MEDLINE data from a preceding block of years (MBD); see SOC for date ranges in the current edition. Separate counts are provided for the frequencies with which the first concept was qualified by different MeSH qualifiers or by no qualifier at all when it co-occurred with the second concept. There are separate entries for each direction of the co-occurrence relationship. The related subheading occurrence information in each entry belongs to the first concept in the entry and is therefore different for each direction of the relationship.

In addition to the specific qualifier information associated with two co-occurring concepts, this element also includes in entries with LQ and LQB values for type of co-occurrence, totals for the number of times each main concept was qualified by a specific subheading or by no subheading.

The AI/RHEUM co-occurrence data represent the co-occurrence of diseases and findings in the AI/RHEUM knowledge base, i.e., the diseases that co-occur with a particular finding and the findings that co-occur with a particular disease. Each disease/finding pair can co-occur only once in the AI/RHEUM knowledge base.

In CCPSS, the co-occurrence data is extracted from patient records and includes problem-problem co-occurrences within a patient record as well as problem-modifier co-occurrences.

Sample records:

C0000294|A0085139|C0002421|A0022413|MBD|L|1|TU=1||.
C0000294|A0085139|C0003968|A0026910|MED|L|1|UR=1||.
C0000294|A0085139|C0006463|A0033415|MBD|L|1|<>=1||.

2.7.1.3.11 "Computable" Hierarchies (File = MRHIER.RRF)

This file contains one row for each hierarchy or context in which each atom appears. If a source vocabulary does not contain hierarchies, its atoms will have no rows in this file. If a source vocabulary is multi-hierarchical (allows the same atom to appear in more than one hierarchy), some of its atoms will have more than one row in this file. MRHIER.RRF provides a complete and compact representation of all hierarchies present in all Metathesaurus source vocabularies. Hierarchical displays can be computed by combining data in this file with data in MRCONSO.RRF. The distance-1 relationships, i.e., immediate parent, immediate child, and sibling relationships, represented in MRHIER.RRF also appear in MRREL.RRF. Most of the hierarchical relationships in MRHIER.RRF (excluding some sibling relationships) also appear in a much larger, pre-computed format in MRCXT.RRF (Section 2.7.1.3.12). NLM plans to phase out MRCXT.RRF (which has reached an unwieldy size) in favor of providing users with tools that generate hierarchical displays based on MRHIER.RRF and MRCONSO.RRF.

Col.	Description
CUI	Unique identifier for concept
AUI	Unique identifier for atom
CXN	Context number (e.g., 1,2,3)
PAUI	Unique identifier of atom's immediate parent within this context
SAB	Source of atom (and therefore of hierarchical context)
RELA	Relationship of atom to its immediate parent
PTR	Path to the top or root of the hierarchical context from this atom, represented as a list of AUIs, separated by periods (.) The first one in the list is top of the hierarchy; the last one in the list is the immediate parent of the atom, which also appears as the value of PAUI.
HCD	Source asserted hierarchical number or code for this atom in this context
CVF	Content view flag
	Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use.

Sample Records

C0001175|A2878223|1|A3316611|SNOMEDCT|isa|A3684559.A2880798.A339606.A3287869.A3316611|||
C0001175|A2878223|2|A3512124|SNOMEDCT|isa|A3684559.A2880798.A3398606.A3287869.A3512124|||
C0001175|A2878223|3|A3696836|SNOMEDCT|isa|A3684559.A2880798.A3398606.A3399957.A3399109.A3144217.A3696836|||
C0001175|A2878223|4|A3512124|SNOMEDCT|isa|A3684559.A2880798.A3398606.A3399957.A3399109.A3512124|||
C0001175|A2878223|5|A3316611|SNOMEDCT|isa|A3684559.A2880798.A3512117.A3082701.A3316611|||
C0001175|A2878223|6|A2888699|SNOMEDCT|isa|A3684559.A2880798.A3512117.A3082701.A3398847.A3398762.A2888699|||
C0001175|A2878223|7|A3316611|SNOMEDCT|isa|A3684559.A2880798.A3512117.A3287869.A3316611|||
C0001175|A2878223|8|A3512124|SNOMEDCT|isa|A3684559.A2880798.A3512117.A3287869.A3512124|||
C0001175|A2988194|1|A2888699|SNOMEDCT|isa|A3684559.A2880798.A3512117.A3082701.A3398847.A3398762.A2888699|||

To find the specific concept names used in a hierarchy, look up the atom identifiers in the AUI and STR data elements in MRCONSO.RRF.

For most source vocabularies, the value of RELA (if present) applies up the hierarchy to the top or root. In other words, it also applies to the relationship between the atom' s parent and the atom' s grandparent, etc. The two exceptions in this version of the Metathesaurus are GO (Gene Ontology) and NIC (Nursing Intervention Classification). Except for GO and NIC atoms, the MRHIER rows for an atom' s ancestors (parent, grandparent, etc.) contain no added information except the source-asserted hierarchical number or code (HCD). If this is not of interest, there may be no reason to find MRHIER rows for an atom' s ancestors.

To find an atom' s siblings in a specific context, find all MRHIER.RRF rows that share its SAB, RELA*, and PTR values.

To find an atom' s children in a specific context, append a period (.) and the atom' s AUI to its PTR and find all MRHIER. RRF rows with its SAB, RELA*, and the expanded PTR.

*The RELA is needed to retrieve correct siblings and children for University of Washington Digital Anatomist (UWDA) hierarchies. Some UWDA atoms appear in multiple hierarchies that are distinguished ONLY by their RELA values.

2.7.1.3.12 Contexts (File = MRCXT.RRF)

This very large file contains pre-computed hierarchical context information (including concept names) intended to facilitate the display of hierarchies present in UMLS source vocabularies. All of the information in this file (plus additional sibling relationships) can be computed by joining the MRHIER.RRF file with MRCONSO.RRF. There can be many rows in this file for each occurrence of an atom in a hierarchy in any of the UMLS source vocabularies - a "context in" this discussion. Many Metathesaurus concepts have many atoms with contexts while others may have none. The number of rows per context differs depending on the number of ancestor, sibling, or child terms an atom has in that context. Because some atoms have multiple contexts in the same source, e.g., MeSH, a context number (CXN - e.g., 1,2,3) is used to identify all members of the same context. The CXNs are not global but are created as required for each atom. Each distinct context for a single atom can be retrieved with a CUI-AUI-SAB-CXN key. The "distance-1 relationships, " i.e., the immediate parent, immediate child, and sibling relationships, represented in MRCXT.RRF, are also present in the MRREL.RRF file.

Col.	Description
CUI	Unique identifier of concept
SUI	Unique identifier for string used in this context
AUI	Unique identifier for atom that has this context
SAB	Source abbreviation. Allowed values appear in MRSAB.RRF and are listed in Appendix B.4
CODE	Unique Identifier or code for string in that source
CXN	The context number (if the atom has multiple contexts)
CXL	Context member label, i.e., ANC for ancestor of this atom, CCP for the atom itself, SIB for sibling of this atom, CHD for child of this atom
RNK	For rows with a CXL value of ANC, the rank of the ancestors (e.g., a value of 1 denotes the most remote ancestor in the hierarchy)
CXS	String or concept name for context member
CUI2	Concept identifier of context member (may be empty if context member is not yet in the Metathesaurus)
AUI2	Atom identifier of context member
HCD	Source hierarchical number or code of context member (if present).
RELA	Additional relationship label providing further categorization of the CXL, if applicable and known. Valid values listed in Appendix B.3.
XC	A plus(+) sign indicates that the CUI2 for this row has children in this context. If this field is empty, the CUI2 does not have children in this context

CVF   Content view flag

           Bit field used to flag rows included in Content View.

Sample Records

C0001175|S1911299|A1855909|ICPC2P|B9001|1|ANC|1|ICPC2-Plus|C1140253|A1861145|||||
C0001175|S1911299|A1855909|ICPC2P|B90001|1|ANC|2|BLOOD/BLOOD FORMING ORGANS/IMMUNE
    MECHANISM|C0847039|A1852564|B||||
C0001175|S1911299|A1855909|ICPC2P|B90001|1|ANC|2|Diagnosis/Diseases Component|C0497531|A0916974|7||||
C0001175|S1911299|A1855909|ICPC2P|B90001|1|ANC|3|HIV-INFECTION|AIDS|C0497169|A1852069|B90||||| C0001175|
S1911299|A1855909|ICPC2P|B90001|1|CCP||Acquired Immune-Deficiency Syndrome|C0001175|A1855909|B90001|||||

2.7.1.3.13 Mappings (File = MRMAP.RRF)

Representations of simple and complex mappings between (1) concept names or (usually) their surrogates (identifiers or codes) from one source vocabulary or from the Metathesaurus and (2) concept names or (usually) their surrogates from another source vocabulary or from the Metathesaurus. This file can accommodate multiple purpose-specific mappings between the same source vocabularies and/or conditional rules for when mappings apply. Source asserted historical mappings (i.e., mappings between obsolete terms/concepts and current ones) are included here.

Col.	Description
MAPSETCUI	Unique identifier for the map set to which this mapping belongs
MAPSETSAB	Source abbreviation for the map set
MAPSUBSETID	Map subset identifier (optional)
MAPRANK	Order in which mappings in a subset should be applied (optional)
MAPID	Mapping ID (ATUI value of XMAP)
MAPSID	Source asserted Mapping ID
FROMID	Mapped_from identifier (source-id assigned by the Metathesaurus as a simple id for what may be a complex expression in FROMEXPR)
FROMSID	Source asserted indentifier for mapped_from
FROMEXPR	Mapped_from expression, which can be a single identifier or concept name or a complex expresion involving multiple identifiers or concept names, Boolean opertors, and/or punctuation
FROMTYPE	Type of mapped_from expression
FROMRULE	Machine processible rule for when the mapped_from is valid (optional)
FROMRES	Restriction on when the mapped_from should be used (optional)
REL	Relationship
RELA	Additional relationship label (optional)
TOID	Mapped_to identifier (target id assigned by the Metathesaurus as a simple id for what may be a complex expression in TOEXPR)
TOSID	Source asserted indentifier for mapped_to
TOEXPR	Mapped_to expression, which can be a single identifier or concept name or a complex expression involving multiple identifiers or concept names, Boolean operators, and/or punctuation
TOTYPE	Type of mapped _to expression
TORULE	Machine processible rule for when the mapped_to is valid (optional)

	TORES	Restriction on when the mapped_to should be used (optional)
	MAPRULE	Machine processible rule for when to apply mapping (optional)
	MAPRES	Human readable rule for use of mapping
	MAPTYPE	Type of mapping
	MAPATN	Row level attribute name associated with this mapping [not yet in use]
	MAPATV	Row level attribute value associated with this mapping [not yet in use]
	CVF	Content view flag

Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use.

Sample Records:
C1306694|MTH|||C0010700||C0010700|CUI|||RB||1552|| <Bladder>/<surgery> | BOOLEAN_EXPRESSION | | | | ATX | | | |
C1306694|MTH|||C0011764||C0011764|CUI|||RO||2201|| <Developmental Disabilities> AND <Writing> | BOOLEAN_EXPRESSION | | | | ATX | | | |

2.7.1.3.14 Simple Mappings (File = MRSMAP.RRF)

A simpler representation of most of the mappings in MRMAP.RRF. This file is provided to serve applications which do not require the full richness of the MRMAP.RRF data structure. It does not include entries for mappings that have MAPSUBSETID and MAPRANK values in MRMAP.RRF.

Col.	Description
MAPSETCUI	Unique identifier for the map set
MAPSETSAB	Source abbreviation for the map set
FROMEXPR	Mapped_from expression
FROMTYPE	Type of mapped_from expression
REL	Relationship
RELA	Additional relationship label
TOEXPR	Mapped_to expression
TOTYPE	Type of mapped_to expression
CVF	Content view flag

Bit field used to flag rows included in Content View. This field is a varchar field to maximize the number of bits available for use.

Sample Records:
C1306694 | MTH | C0009215 |CUI | SY | | <Codeine> AND <Drug Hypersensitivity> | BOOLEAN_EXPRESSION | | |
C1306694 | MTH | C0009216 |CUI | RO | | <Codeine> AND <Drug Tolerance> | BOOLEAN_EXPRESSION | ||

2.7.1.3.15 Source Information (File=MRSAB.RRF)

The UMLS Metathesaurus has "versionless" or "root" Source Abbreviations (SABs) in the data files. MRSAB.RRF connects the "root" SAB to fully specified version information for the current release. For example, the released SAB for MeSH is now simply "MSH". In MRSAB.RRF, you will see a current versioned SAB, e.g., MSH2003_2002_10_24. MRSAB.RRF allows

all other Metathesaurus files to use versionless source abbreviations, so that all rows with no data change between versions remain unchanged. MetamorphoSys can produce files with either the root or versioned SABs so that either form can be available in custom subsets of the Metathesaurus.

There is one row in this file for every version of every source in the current Metathesaurus; eventually there will also be historical information with a row for each version of each source that has appeared in any Metathesaurus release. Note that the field CURVER has the value 'Y' to identify the version in this Metathesaurus release. Future releases of MRSAB.RRF will also contain historical version information in rows with CURVER value 'N'.

The structure of MRSAB.RRF is as follows:

Field	Full Name	Description
VCUI	CUI	CUI of the versioned SRC concept for a source
RCUI	Root CUI	CUI of the root SRC concept for a source
VSAB	Versioned Source Abbreviation	The versioned source abbreviation for a source, e.g., MSH2003_2002_10_24
RSAB	Root Source Abbreviation	The root source abbreviationm, for a source e.g MSH
SON	Official Name	The official name for a source
SF	Source Family	The Source Family for a source
SVER	Version	The source version, e.g., 2001
VSTART	Meta Start Date	The date a source became active, e.g., 2001_04_03
VEND	Meta End Date	The date a source ceased to be active, e.g., 2001_05_10
IMETA	Meta Insert Version	The version of the Metathesaurus in which a source first appeared, e.g., 2001AB
RMETA	Meta Remove Version	The version of the Metathesaurus in which the source last appeared, e.g., 2001AC
SLC	Source License Contact	The source license contact information
SCC	Source Content Contact	The source content contact information
SRL	Source Restriction Level	0,1,2,3,4 - explained in the License Agreement.
TFR	Term Frequency	The number of terms for this source in MRCONSO.RRF, e.g., 12343
CFR	CUI Frequency	The number of CUIs associated with this source, e.g., 10234
CXTY	Context Type	The type of contexts for this source. Values are FULL, FULL-MULTIPLE, FULL-NOSIB, FULL-NOSIB-MULTIPLE, FULL-MULTIPLE-NOSIB-RELA, null.
TTYL	Term Type List	Term type list from source, e.g., MH,EN,PM,TQ
ATNL	Attribute Name List	The attribute name list (from MRSAT.RRF), e.g., MUI,RN,TH,...
LAT	Language	The language of the terms in the source
CENC	Character Encoding	All UMLS content is provided in Unicode, encoded in UTF-8. MetamorphoSys will allow exclusion of extended characters with some loss of information. Transliteration to other character encodings is possible but not supported buy NLM; for further information, see http://www.unicode.org.
CURVER	Current Version	A Y or N flag indicating whether or not this row corresponds to the current version of the named source

SABIN	Source in Subset	A Y or N flag indicating whether or not this row is represented in the current MetamorphoSys subset. Initially always Y where CURVER is Y, but later is recomputed by MetamorphoSys.
SSN	Source short name	The short name of a source as used by the NLM Knowledge Source Server.
SCIT	Source citation	Citation information for a source. This is intented to replace the SOS attributes in the SRC concepts.

Sources with contexts have "full" contexts, i.e., all levels of terms may have Ancestors, Parents, Children and Siblings. A full context may also be further designated as Multiple, Nosib (No siblings) or both Multiple and Nosib.
"Multiple" indicates that a single concept in this source may have multiple hierarchical positions.

No siblings (Nosib) indicates that siblings have not been computed for this source.
Appendix B.4, Source Vocabularies, lists each source in the Metathesaurus and includes information about the type of context, if any, for each source.

Sample Records:
C1371270|C1140284|RXNORM_04AB|RXNORM|RXNORM Project, META2004AB | RXNORM | 04AB | 2004_05_17 | | 2004AB | | Stuart Nelson, M.D., Head, MeSH Section; e-mail: nelson@nlm.nih.gov | Stuart Nelson, M.D., Head, MeSH Section; e-mail: nelson@nlm.nih.gov | 0 | 138005 | 110403 | | BN,IN,OBD,OCD,SBD,SBDF,SCD,SCDC,SCDF,SY | ORIG_CODE,ORIG_SOURCE | ENG | UTF-8 | Y | Y | RxNorm work done by the National Library of Medicine|RxNorm work done by NLM. Bethesda (MD): National Library of Medicine, META2004AB release. | |

2.7.1.3.16 Concept Name Ranking (File=MRRANK.RRF)

There is exactly one row for each concept name type from each Metathesaurus source vocabulary (each SAB-TTY combination). The RANK and SUPPRES values in the distributed file are those used in Metathesaurus production. Users are free to change these values to suit their needs and preferences, then change the naming precedence and suppressibility by using MetamorphoSys to create a customized Metathesaurus.

Col.	Description
RANK	Numeric order of precedence, higher value wins
SAB	Abbreviation for source vocabulary
TTY	Abbreviation for concept name type in source vocabulary
SUPPRESS	NLM-recommended Source and Term Type (SAB/TTY) Suppressiblity. Values = Y or N. Indicates the suppressible status of all atoms (names) with this Source and Term Type (SAB/TTY). Note that changes made in MetamorphoSys at the Suppressible tab are recorded in your configuration file. Status E does not occur here, as it is assigned only to individual cases such as the names (atoms) in MRCONSO.RRF. See also SUPPRESS in MRCONSO.RRF and MRDEF.RRF and MRREL.RRF.

Sample Records

0210|AIR|SY|N|
0209|ULT|PT|N|
0208|CPT|PT|N|

2.7.1.3.17 Ambiguous Term Identifiers (File = AMBIGLUI.RRF)

In the instance that a Lexical Unique Identifier (LUI) is linked to multiple Concept Unique Identifiers (CUIs), there is one

row in this table for each LUI-CUIs pair. This file identifies those lexical variant classes which have multiple meanings in the Metathesaurus.

In the Metathesaurus, the LUI links all strings within the English language that are identified as lexical variants of each other by the luinorm program found in the UMLS SPECIALIST Lexicon and Tools. LUIs are assigned irrespective of the meaning of each string. This table may be useful to system developers who wish to make use of the lexical programs in their applications to identify and disambiguate ambiguous terms.

Col.	Description
LUI	Lexical Unique Identifier
CUI	Concept Unique Identifier

Sample Records

L0000003|C0010504|
L0000003|C0917995|
L0000032|C0010206|

2.7.1.3.18 Ambiguous String Identifiers (File=AMBIGSUI.RRF)

In the instance that a String Unique Identifier (SUI) is linked to multiple Concept Unique Identifiers (CUIs), there is one row in this table for each SUI-CUIs pair.

This file resides in the META directory. In the Metathesaurus, there is only one SUI for each unique string within each language, even if the string has multiple meanings. This table is only of interest to system developers who make use of the SUI in their applications or in local data files.

Col.	Description
SUI	String Unique Identifier
CUI	Concept Unique Identifier

Sample Records

S0063890|C0026667|
S0063890|C1135584|
S0229413|C0008802|

2.7.1.3.19 Metathesaurus Change Files

There are six files or relations that identify key differences between entries in the previous and the current edition of the Metathesaurus. Developers can use these special files to determine whether there have been changes that affect their applications.

The usefulness of individual files will depend on how data from the Metathesaurus have been linked or incorporated in a particular application.

Each relation or named table of data has a fixed number of columns and variable number of rows. A column is a sequence of all the values in a given data element. A row contains the values for two or more data elements for one entry. The values for the different data elements in the row are separated by vertical bars (|). Each row ends with a vertical bar and line termination.

2.7.1.3.19.1 Deleted Concepts (File=CHANGE/DELETEDCUI.RRF)

There is exactly one row in this table for each reviewed concept that was present in the previous Metathesaurus and is not present in the current Metathesaurus.

Cols.
PCUI Concept Unique Identifier in the previous Metathesaurus
PSTR Preferred name of this concept in the previous Metathesaurus

2.7.1.3.19.2 Merged Concepts (File=CHANGE/MERGEDCUI.RRF)

There is exactly one row in this table for each released concept in the previous Metathesaurus (CUI1) that was merged into another released concept from the previous Metathesaurus (CUI2). When this merge occurs, the first CUI (CUI1) was retired; this table shows the CUI (CUI2) for the merged concept in this Metathesaurus.

Entries in this file represent concepts pairs that were considered to have different meanings in the previous edition, but which are now identified as synonyms

Cols.
PCUI1 Concept Unique Identifier in the previous Metathesaurus
CUI Concept Unique Identifier in this Metathesaurus in format C#######

2.7.1.3.19.3 Deleted Terms (File=CHANGE/DELETEDLUI.RRF)

There is exactly one row in this table for each Lexical Unique Identifier (LUI) that appeared in the previous Metathesaurus, but does not appear in this Metathesaurus.

Metathesaurus Lexical Unique Identifiers (LUIs) are assigned by the luinorm program, part of LVG program in the UMLS SPECIALIST Lexicon and Tools; see Section 4.

These entries represent the cases where LUIs identified by the previous release's luinorm program, when used to identify lexical variants in the previous Metathesaurus, are no longer found with this release's luinorm on this release's Metathesaurus. This does not necessarily imply the deletion of a string or a concept from the Metathesaurus.

Cols.
PLUI Lexical Unique Identifier in the previous Metathesaurus
PSTR Preferred Name of Term in the previous Metathesaurus

2.7.1.3.19.4 Merged Terms (File=CHANGE/MERGEDLUI.RRF)

There is exactly one row in this file for each case in which strings had different Lexical Unique Identifiers (LUIs) in the previous Metathesaurus yet share the same LUI in this Metathesaurus; a LUI present in the previous Metathesaurus is therefore absent from this Metathesaurus.

Metathesaurus Lexical Unique Identifiers (LUIs) are assigned by the luinorm program, part of the LVG program in the UMLS SPECIALIST Lexicon and Tools; see Section 4.

These entries represent the cases where separate lexical variants as identified by the previous release's luinorm program version are a single lexical variant as identified by this release's luinorm.

Cols.
PLUI Lexical Unique Identifier in the previous Metathesaurus but not present in this Metathesaurus
LUI Lexical Unique Identifier into which it was merged in this Metathesaurus

2.7.1.3.19.5 Deleted Strings (File=CHANGE/DELETEDSUI.RRF)

There is exactly one row in this file for each string in each language that was present in a entry in the previous Metathesaurus and does not appear in this Metathesaurus.

Note that this does not necessarily imply the deletion of a term (LUI) or a concept (CUI) from the Metathesaurus. A string deleted in one language may still appear in the Metathesaurus in another language.

Cols.
PSUI String Unique Identifier in previous Metathesaurus that is not present in this Metathesaurus
PSTR Preferred name of term in previous Metathesaurus that is not present in this Metathesaurus

2.7.1.3.19.6 Retired CUI Mapping (File=MRCUI.RRF)

There are one or more rows in this file for each Concept Unique Identifier (CUI) that existed in any prior release but is not present in the current release. The file includes mappings to current CUIs as synonymous or to one or more related current CUI where possible. If a synonymous mapping cannot be found, other relationships between the CUIs can be created. These relationships can be Broader (RB), Narrower (RN), Other Related (RO), Deleted (DEL) or Removed from Subset (SUBX). Rows with the SUBX relationship are added to MRCUI by MetamorphoSys for each CUI that met the exclusion criteria and was consequently removed from the subset. Some CUIs may be mapped to more than one other CUI using these relationships.

CUIs may be retired when (1) two released concepts are found to be synonyms and so are merged, retiring one CUI; (2) when the concept no longer appears in any source vocabulary and is not 'rescued' by NLM; or (3) where the concept is an acknowledged error in a source vocabulary or determined to be a Metathesaurus production error.

See Sections 2.7.1.3.19, 1 through 5 for filesof changes from the last release only, without mappings.

Col.	Description
CUI1	Unique identifier for first concept -- Retired CUI - was present in some prior release, but is currently missing
VER	The last release version in which CUI1 was a valid CUI
REL	Relationship
RELA	Relationship attribute
MAPREASON	Reason for mapping
CUI2	Unique identifier for second concept -- The current CUI that CUI1 most closely maps to.
MAPIN	Is this map in current subset? Values of Y or N or null. MetamorphoSys generates the Y or N to indicate whether the CUI2 concept is or is not present in the subset. The null value is for rows where the CUI1 was not present to begin with (i.e., REL=DEL).

Sample Records:
C1313903|2004AA|SY|||C0525045|Y|
C1313909|2004AA|RO|||C0476661|Y|
C1321833|2004AA|DEL|||||
C1382264|2004AB|SY|||C0993613|Y|
C1382494|2004AB|DEL|||||

2.7.1.3.19.7 AUI Movements and Deletions (File=MRAUI.RRF)

This file records the movement of Atom Unique Identifiers (AUIs) from one concept (CUI1) in a previous version of the UMLS Metathesaurus (VER) to another concept (CUI2).

Col.	Description
AUI1	Atom Unique identifier
CUI1	Concept Unique identifier
VER	version in which this change to the AUI occurred
REL	Relationship
RELA	Relationship attribute
MAPREASON	Reason for mapping
AUI2	Unique identifier for second atom
CUI2	Unique identifier for second concept -- The current CUI that CUI1 most closely maps to.
MAPIN	Mapping in current subset: is AUI2 in current subset? Values of Y or N or null.

Sample Records:

A0000039|C0236824|2004AC|||move|A0000039|C1411876|Y|
A0000077|C0003477|2005AB|||move|A0000077|C1510447|Y|
A8177040|C1237728|2005AB|||move|A8177040|C1237732|Y|

2.7.1.3.20 Word Index (File = MRXW_BAQ.RRF, MRXW_DAN.RRF, MRXW_DUT.RRF, MRXW_ENG.RRF, MRXW_FIN.RRF, MRXW_FRE.RRF, MRXW_GER.RRF, MRXW_HEB.RRF, MRXW_HUN.RRF, MRXW_ITA. RRF, MRXW_NOR.RRF, MRXW_POR.RRF, MRXW_RUS.RRF, MRXW_SPA.RRF, MRXW_SWE.RRF)

There is one row in these tables for each word found in each unique Metathesaurus string (ignoring upper-lower case). All Metathesaurus entries have entries in the word index. The entries are sorted in ASCII order.

Col.	Description
LAT	Abbreviation of language of the string in which the word appears
WD	Word in lowercase
CUI	Concept identifier
LUI	Term identifier
SUI	String identifier

Sample Records from MRXW_ENG.RRF

ENG|anaemia|C0002871|L0280031|S0352688|
ENG|anemia|C0002871|L0002871|S0013742|
ENG|anemias|C0002871|L0002871|S0013787|
ENG|blood|C0002871|L0376533|S0500659|

Sample Records from MRXW_FRE.RRF

FRE|ANEMIE|C0002871|L0162748|S0227229|

2.7.1.3.21 Normalized Word Index (File=MRXNW_ENG.RRF)

There is one row in this table for each normalized word found in each unique English-language Metathesaurus string. All English-language Metathesaurus entries have entries in the normalized word index. There are no normalized string indexes for other languages in this edition of the Metathesaurus.

Col.	Description
LAT	Abbreviation of language of the string in which the word appears (always ENG in this edition of the Metathesaurus)
NWD	Normalized word in lowercase (described in Section 2.6.2.1)
CUI	Concept identifier
LUI	Term identifier
SUI	String identifier

Sample Records

ENG|anaemia|C0002871|L0280031|S0352688|
ENG|anemia|C0002871|L0002871|S0013742|
ENG|anemia|C0002871|L0002871|S0013787|
ENG|blood|C0002871|L0376533|S0500659|

2.7.1.3.22 Normalized String Index (File=MRXNS_ENG.RRF)

There is one row in this table for each normalized string found in each unique English-language Metathesaurus string (ignoring upper-lower case). All English-language Metathesaurus entries have entries in the normalized string index. There are no normalized word indexes for other languages in this edition of the Metathesaurus.

Col.	Description
LAT	Abbreviation of language of the string (always ENG in this edition of the Metathesaurus)
NSTR	Normalized string in lowercase (described in Section 2.6.3.1)
CUI	Concept identifier
LUI	Term identifier
SUI	String identifier

Sample Records

ENG|anaemia|C0002871|L0280031|S0352688|
ENG|anaemia unspecified|C0002871|L0696700|S0803315|
ENG|anemia|C0002871|L0002871|S0013787|

2.7.2 METATHESAURUS ORIGINAL RELEASE FORMAT (ORF)

Note: The preferred and more complete format is described above in Section 2.7.1, the Metathesaurus Rich Release Format (RRF).

All files except MRRANK are sorted by row.

2.7.2.1. Data Files

The data in each Metathesaurus entry may be represented in more than 20 different "relations" or files. These files correspond

to the four logical groups of data elements described in Section 2.2 - 2.5 and the indexes described in Section 2.6 as follows:

Metathesaurus Concept Names and their sources (2.7.2.2) = MRCON, MRSO
Attributes (2.7.2.3) = MRSAT, MRDEF, MRSTY
Relationships between Different Concept Names (2.7.2.4) = MRREL, MRCOC, MRATX, MRCXT
Data about the Metathesaurus (2.7.2.5)=MRSAB, MRRANK, AMBIG.LUI, AMBIG.SUI, DELETED.CUI, MERGED.CUI, DELETED.LUI, MERGED.LUI, DELETED.SUI, MRCUI
Indexes (2.7.2.6) = MRXW.BAQ, MRXW.DAN, MRXW.DUT, MRXW.ENG, MRXW.FIN, MRXW.FRE, MRXW.GER, MRXW.HEB, MRXW.HUN, MRXW.ITA, MRXW.NOR, MRXW.POR, MRXW.RUS,MRXW.SPA, MRXW.SWE, MRXNW.ENG, MRXNS.ENG

The AMBIG* files now provide a convenient way to identify all Metathesaurus terms and strings that have more than one meaning in Metathesaurus source vocabularies.

2.7.2.2 Columns and Rows

Each relation or named table of data values has by definition a fixed number of columns; the number of rows depends on the content of a particular version of the Metathesaurus.

A column is a sequence of all the values in a given data element or logical subelement. In general, columns for longer variable length data elements will appear to the right of columns for shorter and/or fixed length data elements. The information for all columns in the ORF files is described in Appendix B.1.2, ORF Columns or Data Elements

A row contains the values for one or more data elements or logical subelements for one Metathesaurus entry. Depending on the nature of the data elements involved, each Metathesaurus entry may have one or more rows in a given file. The values for the different data elements or logical subelements represented in the row are separated by vertical bars (|). If an optional element is blank, the vertical bars are still used to maintain the correct positioning of the subsequent elements. Each row is terminated by a vertical bar and line termination.

2.7.2.3 Descriptions of Each File

The descriptions of the files appear in the following order:
     a.) Key data about the Metathesaurus: Files, Columns or data elements,
     b.) Concept names and their vocabulary sources,
     c.) Attributes
     d.) Relationships
     e.) Other data about the Metathesaurus
     f.) Indexes

2.7.2.3.1 Files (File = MRFILES)

There is exactly one row in this file for each physical segment of the files in the relational format. The columns or data elements in the file are:

Col.	Description
FIL	Physical FILENAME
DES	Descriptive Name
FMT	Comma separated list of COL, in order
CLS	# of COLUMNS
RWS	# of ROWS

BTS   Size in bytes in this format (ISO/PC or Unix)

Sample Records

MRATX|Associated Expressions|CUI,SAB,REL,ATX|4|7295|442571|
MRCOC|Co-occurring Concepts|CUI1,CUI2,SAB,COT,COF,COA|6|9061980|343331578|
MRCOLS|Attribute Relation|COL,DES,REF,MIN,AV,MAX,FIL, DTY|8|115|5728|

2.7.2.3.2 Data Elements (File = MRCOLS)

There is exactly one row in this file for each column or data element in each file in the relational format.

Col.	Description
COL	Column or data element name
DES	Descriptive Name
REF	Documentation Section Number
MIN	Minimum Length, Characters
AV	Average Length
MAX	Maximum Length, Characters
FIL	Physical FILENAME in which this field occurs
DTY	SQL-92 data type for this column

Sample Records

ATN|Attribute name||2|3.15|7|MRSAT|varchar(20)|
ATV|Attribute value||1|9.71|3634|MRSAT|varchar(4000)|
ATX|Associated expression||5|35.89|242|MRATX|varchar(300)|

2.7.2.3.3 Concept Names (File = MRCON)

There is exactly one row in this file for each meaning of each unique string in the Metathesaurus, i.e., there is exactly one row for each unique CUI-SUI combination in the Metathesaurus. Any difference in upper-lower case, word order, etc. creates a different unique string.

Col.	Description
CUI	Unique identifier for concept
LAT	Language of Term
TS	Term status
LUI	Unique identifier for term
STT	String type
SUI	Unique identifier for string
STR	String
LRL	Least Restriction Level

Sample Records

C0002871|ENG|P|L0002871|PF|S0013742|Anemia|0|
C0002871|ENG|P|L0002871|VP|S0013787|Anemias|0|
C0002871|ENG|P|L0002871|VC|S0352787|ANEMIA|0|
C0002871|ENG|P|L0002871|VC|S0414880|anemia|0|
C0002871|ENG|P|L0002871|VO|S0470197|Anemia, NOS|3|
C0002871|ENG|S|L0280031|PF|S0803242|Anaemia|3|

2.7.2.3.4 Vocabulary Sources (File = MRSO)

The vocabulary source(s) for a concept, term, and string.

There is exactly one row in this file for each source of each string in the Metathesaurus. All Metathesaurus concepts have entries in this file.

Col.	Description
CUI	Unique identifier for concept
LUI	Unique identifier for term
SUI	Unique identifier for string
SAB	Source abbreviation. Allowed values are listed in Appendix B, Section B.2
TTY	Term type in that source. Allowed values are listed in Appendix B., Section B.4.
CODE	Unique Identifier or code for string in that source.
SRL	Source Restriction Level

Sample Records

C0002871|L0002871|S0013742|CCS|MD|4.1|0|
C0002871|L0002871|S0013742|ICPCPAE|PT|B82005|3|
C0002871|L0002871|S0013742|LCH|PT|U000235|0|
C0002871|L0002871|S0013742|MSH|MH|D000740|0|
C0002871|L0002871|S0013742|MTH|PT|U000161|0|
C0002871|L0002871|S0013742|MTH|PT|U000164|0|
C0002871|L0002871|S0013742|PSY|PT|02450|3|
C0002871|L0002871|S0013742|RCDAE|PT|XM05A|3|

The information in MRSO can be used in combination with MRCON to determine whether a particular concept, name, or code is present in a particular source, and in what form it appears.

Note: In the RRF, the concept name and vocabulary source information appear in a single file, MRCONSO.RRF.

2.7.2.3.5 Simple Concept and String Attributes (File = MRSAT)

There is exactly one row in this table for each concept, term and string attribute that does not have a sub-element structure. All Metathesaurus concepts have entries in this file.

Col.	Description
CUI	Unique identifier for concept
LUI	Unique identifier for term (optional)

	SUI	Unique identifier for string (optional)
	CODE	Unique identifier or code for entry in the source of the attribute, e.g., for all attributes derived from MeSH, the MeSH unique identifier (optional).
	ATN	Attribute name. Possible values are all described in Appendix B, Section B.1.2.
	SAB	Abbreviation of the source of the attribute. Allowed values are listed in Appendix B, Section B.2.)
	ATV	Attribute value described under specific attribute name in Appendix B, Section B.1.2. A few attribute values exceed 1,000 characters.

Sample Records

C0002871|L0002871|S0013742|D000740|MMR|MSH|19960610|
C0002871|L0002871|S0013742|D000740|MN|MSH|C15.378.71|
C0002871|L0002871|S0013742|D000740|TH|MSH|POPLINE (1994)|
C0002871|L0002871|S0414880|208/04453|SOS|PDQ|secondary related condition|
C0002871|L0002871|S0470197|DC-10010|SIC|SNMI|285.9|

2.7.2.3.6 Definitions (File = MRDEF)

There is exactly one row in this file for each definition in the Metathesaurus. A few definitions approach 3,000 characters in length.

Col.	**Description**
CUI	Unique identifier for concept
SAB	Abbreviation of the source of the definition
DEF	Definition

Sample Records

C0002871|MSH|A reduction in the number of circulating erythrocytes or in the quantity of hemoglobin.|

2.7.2.3.7 Semantic Types (File = MRSTY)

There is exactly one row in this file for each semantic type assigned to each concept. All Metathesaurus concepts have at least one entry in this file. Many have more than one entry.

Col.	**Description**
CUI	Unique identifier of concept
TUI	Unique identifier of Semantic type
STY	Semantic type. The valid values are defined in the Semantic Network.

Sample Record

C0002871|T047|Disease or Syndrome|

2.7.2.3.8 Locators (File = MRLO)

This file has been deleted from the Metathesaurus effective with the 2004AB release. Some of the information was outdated, some was duplicative of information contained in other Metathesaurus files, and some was easily obtained from other publicly available sources, e.g., PubMed.

2.7.2.3.9 Related Concepts (File = MRREL)

There is one row in this table for each relationship between Metathesaurus concepts known to the Metathesaurus, with the following exceptions found in other files: co-occurrences found in MRCOC, and Associated Expressions found in MRATX.

Note that for asymmetrical relationships there is one row for each direction of the relationship. Note also the direction of REL - the relationship which the SECOND concept (with Concept Unique Identifier CUI2) HAS TO the FIRST concept (with Concept Unique Identifier CUI1).

Col.	Description
CUI1	Unique identifier of first concept
REL	Relationship of SECOND to first concept
CUI2	Unique identifier of second concept
RELA	Relationship attribute
SAB	Abbreviation of the source of relationship
SL	Source of relationship labels
MG	Machine-generated and unverified indicator (optional) G indicates 'machine generated'

Sample Records

```
C0002871|CHD|C0002891|isa|MSH|MSH||

  Anemia, Neonatal (C0002891)

    has CHILD REL and isa RELA

      to Anemia (C0002871)


C0002871|RB|C0221016||MTH|MTH||

  [Red blood cell disorder, NOS (C0221016)

    has broader REL

      to Anemia (C0002871)]

C0002871|RL|C0002886|mapped_to|SNMI|SNMI||

  [Anemia, Macrocytic (C0002886)

    has like relationship

      to Anemia (C0002871)]
```

```
C0002871|RO|C0002886|clinically_associated_with|CCPSS|CCPSS||

   [Megaloblastic anemia due to folate deficiency, NOS (C0151482)

     has clinically_associated_with relationship

       to Anemia (C0002871)]
```

2.7.2.3.10 Co-occurring Concepts (File = MRCOC)

There are two rows in this table for each pair of concepts that co-occur in each information source represented one for each direction of the relationship. (Note that the COA data may be different for each direction of the relationship). Many Metathesaurus concepts have no entries in this file. Due to the very large number of co-occurrence relationships, they are distributed in a separate file.

Col.	Description
CUI1	Unique identifier of first concept
CUI2	Unique identifier of second concept Note: Where COT is MeSH topical qualifier (LQ) and CUI2 is not present, the count of citations of CUI1 with no MeSH qualifiers is reported.
SOC	Abbreviation of the Source of co-occurrence information if applicable
COT	Type of co-occurrence
COF	Frequency of co-occurrence, if applicable
COA	Attributes of co-occurrence, if applicable

Sample Records

```
C0002871||MED|LQ|1||
C0002871|C0000530|MBD|L|2|CI=1,EN=1,ME=1,PA=1|
C0002871|C0000727|MBD|L|1|BL=1,ET=1|
C0002871|C0000737|MBD|L|1|ET=1|
C0002871|C0000772|MBD|L|2|CN=2|
```

Co-occurrences are concepts that occur together in the same "entries" in some information source. The relationships represented here are obtained from machine-manipulation of the information source. Co-occurrence relationships may exist between similar concepts (e.g., "Atrial Fibrillation" and "Arrhythmia") or between very different concepts that nevertheless have some important connection in the field of biomedicine (e.g., "Atrial Fibrillation" and "Digoxin"), or between a primary concept and a qualifier e.g., "Lithotripsy" and "instrumentation". A co-occurrence relationship can exist between two concepts that have no other apparent relationship, although the frequency of such co-occurrences will be small.

In the current Metathesaurus, there are three sources of co-occurrence data: MEDLINE, AI/RHEUM, and CCPSS. From MEDLINE, co-occurrence data was computed for concepts that were designated as principal or main points in the same journal article i.e., the co-occurrence counts do not include articles in which either or both of the concepts were present and indexed in MEDLINE but not designated as main points. (A concept is considered to be a main point if the * is attached to the main heading or any of its subheadings.)

Two overall frequencies of MEDLINE co-occurrence are provided: one for recent MEDLINE data (MED) and one for MEDLINE data from a preceding block of years (MBD); see SOC for date ranges in the current edition. Separate counts are provided for the frequencies with which the first concept was qualified by different MeSH qualifiers or by no qualifier at all when it co-occurred with the second concept. There are separate entries for each direction of the co-occurrence relationship.

The related subheading occurrence information in each entry belongs to the first concept in the entry and is therefore different for each direction of the relationship.

In addition to the specific qualifier information associated with two co-occurring concepts, this element also includes in entries with LQ and LQB values for type of co-occurrence, totals for the number of times each main concept was qualified by a specific subheading or by no subheading.

The AI/RHEUM co-occurrence data represent the co-occurrence of diseases and findings in the AI/RHEUM knowledge base, i.e., the diseases that co-occur with a particular finding and the findings that co-occur with a particular disease. Each disease/finding pair can co-occur only once in the AI/RHEUM knowledge base.

In CCPSS, the co-occurrence data is extracted from patient records and includes problem-problem co-occurrences within a patient record as well as problem-modifier co-occurrences.

2.7.2.3.11 Concept contexts (File = MRCXT)

There are rows in this file for each occurrence of a concept in a hierarchy in any of the UMLS source vocabularies - a "context" in this discussion. Many Metathesaurus concepts have multiple contexts while others may have none. The number of rows per context differs depending on the number of ancestor, sibling, or child terms the concept has in that context. Because some concepts have multiple contexts in the same source (e.g., MeSH), a context number (CXN - e.g., 1, 2, 3) is used to identify all members of the same context. The CXNs are not global but are created as required for each concept. Since some concepts have multiple contexts in the same vocabulary with the same SUI, each distinct context can be retrieved with a CUI-SUI-SAB-CXN key. The "distance-1 relationships," i.e., the immediate parent, immediate child, and sibling relationships, represented in this file are also present in the MRREL file.

(Note: The RELA was incorrectly called REL in versions before 2001.)

Col.	Description
CUI	Unique identifier of concept
SUI	Unique identifier for string used in this context
SAB	Source abbreviation. Allowed values are listed in Appendix B.4
CODE	Unique Identifier or code for string in that source.
CXN	The context number (to distinguish multiple contexts in the same source with the same SUI).
CXL	Context member label, i.e., ANC for ancestor of this concept, CCP for concept, SIB for sibling of this concept, CHD for child of this concept.
RNK	For rows with a CXL value of ANC, the rank of the ancestors (e.g., a value of 1 denotes the most remote ancestor in the hierarchy)
CXS	String for context member.
CUI2	Unique concept identifier of context member (may be empty if context member is not yet in the Metathesaurus).
HCD	Hierarchical number or code of context member in this source (optional).
RELA	Relationship attribute providing further categorization of the CXL, if applicable and known. Allowed values are listed in Appendix B.3.
XC	A plus(+) sign indicates that the CUI2 for this row has children in this context. If this field is empty, the CUI2 does not have children in this context.

Sample Records

C0002871|S0013742|MSH|D000740|1|ANC|1|MeSH|C0220876||||

C0002871|S0013742|MSH|D000740|1|ANC|2|Diseases (MeSH Category)|C0012674|C|||
C0002871|S0013742|MSH|D000740|1|ANC|3|Hemic and Lymphatic Diseases|C0018981|C15|||
C0002871|S0013742|MSH|D000740|1|ANC|4|Hematologic Diseases|C0018939|C15.378|isa||
C0002871|S0013742|MSH|D000740|1|CCP||Anemia|C0002871|C15.378.71|isa|+|
C0002871|S0013742|MSH|D000740|1|CHD||Anemia, Aplastic|C0002874|C15.378.71.85|isa|+|
C0002871|S0013742|MSH|D000740|1|SIB||Blood Protein Disorders|C0005830|C15.378.147|isa|+|
C0002871|S0013742|MSH|D000740|1|CHD||Anemia, Hemolytic|C0002878|C15.378.71.141|isa|+|

2.7.2.3.12 Associated Expressions (File = MRATX)

There is one row in this table for each vocabulary expression (i.e., combination of terms from a specific Metathesaurus source vocabulary) identified as having a relationship to a concept in the Metathesaurus. The majority of Metathesaurus entries have no entries in this table.

Col.	Description
CUI	Unique identifier of concept to which the expression is related
SAB	Abbreviation of source of terms in expression. Allowed values are listed in Appendix B, Section B.1)
REL	Relationship of meaning of expression to main concept
ATX	Associated expression

Sample Records

C0001207|MSH|S|<Acromegaly> AND <Gigantism>|
C0001296|LCH|U|<Insurance>/<Statistics>|
C0001355|MSH|S|<Kidney Failure, Acute> AND <Kidney Papillary Necrosis>|

2.7.2.3.13 Source Information (File=MRSAB)

The UMLS Metathesaurus has "versionless" or "root" Source Abbreviations (SABs) in the data files. MRSAB connects the "root" SAB to fully specified version information for the current release. For example, the released SAB for MeSH is now simply "MSH". In MRSAB, you will find the current versioned SAB, e.g., MSH2003_2002_10_24. MetamorphoSys can produce files with either the root or versioned SABs so that either form can be utilized by a user.

There is one row in this file for every version of every source in the current Metathesaurus; when complete, there will also be historical information with a row for each version of each source that has appeared in any Metathesaurus release. Note that the field CURVER has the value 'Y' to identify the version in this Metathesaurus release. Future releases of MRSAB will also contain historical version information in rows with CURVER value 'N'.

MRSAB allows all other Metathesaurus files to use versionless source abbreviations, so that rows with no data change between versions also remain unchanged.

The full structure of MRSAB is as follows:

Field	Full Name	Description
VCUI	CUI	CUI of the versioned SRC concept for a source
RCUI	Root CUI	CUI of the root SRC concept for a source
VSAB	Versioned Source Abbreviation	The versioned source abbreviation for a source e.g. MSH2003_2002_10_24
RSAB	Root Source Abbreviation	The root source abbreviation for a source e.g MSH
SON	Official Name	The official name for a source

SF	Source Family	The Source Family for a source
SVER	Version	The source version e.g. 2001
VSTART	Valid Start Date For A Source	Source's start date for valid use, e.g. 2004_04_03
VEND	Valid End Date For A Source	Source's end date for valid use, e.g. 2003_05_10
IMETA	Meta Insert Version	The version of the Metathesaurus a source first appeared, e.g.2001AB
RMETA	Meta Remove Version	The version of the Metathesaurus a source was removed, e.g.2001AC
SLC	Source License Contact	The source license contact information
SCC	Source Content Contact	The source content contact information
SRL	Source Restriction Level	0,1,2,3
TFR	Term Frequency	The number of terms for this source in MRCON/MRSO, e.g., 12343
CFR	CUI Frequency	The number of CUIs associated with this source, e.g. 10234
CXTY	Context Type	The type of context (per section 2.3.2) from the UMLS documentation
TTYL	Term Type List	Term type list from source , e.g. MH,EN,PM,TQ
ATNL	Attribute Name List	The attribute name list (from MRSAT), e.g., MUI,RN,TH,...
LAT	Language	The language of the source
CENC	Character Encoding	All UMLS content is provided in Unicode, encoded in UTF-8. MetamorphoSys will allow exclusion of extended characters with some loss of information. Transliteration to other character encodings is possible but not supported buy NLM; for further information, see http://www.unicode.org.
CURVER	Current Version	A Y or N flag indicating whether or not this row corresponds to the current version of the named source
SABIN	Source in Subset	A Y or N flag indicating whether or not this row is represented in the current MetamorphoSys subset. Initially always Y where CURVER is Y, but later is recomputed by MetamorphoSys.

Sources with contexts have "full" contexts, i.e., all levels of terms may have Ancestors, Parents, Children and Siblings. A full context may also be further designated as Multiple, Nosib (No siblings) or both Multiple and Nosib.
"Multiple" indicates that a single concept in this source may have multiple hierarchical positions.

No siblings (Nosib) indicates that siblings have not been computed for this source.
Appendix B.4, Source Vocabularies, lists each source in the Metathesaurus and includes information about the type of context, if any, for each source.

Sample Record:

C1371270|C1140284|RXNORM_04AB|RXNORM|RXNORM Project, META2004AB | RXNORM | 04AB | 2004_05_17 | | 2004AB | | Stuart Nelson, M.D., Head, MeSH Section; e-mail: nelson@nlm.nih.gov | Stuart Nelson, M.D., Head, MeSH Section; e-mail: nelson@nlm.nih.gov | 0 138005 | 110403 | | BN,IN,OBD,OCD,SBD,SBDF,SCD,SCDC,SCDF,SY | ORIG_CODE,ORIG_SOURCE | ENG | UTF-8| Y | Y |

2.7.2.3.14 Concept Name Ranking (File=MRRANK)

There is exactly one row for each concept name type from each Metathesaurus source vocabulary (each SAB-TTY

combination). The RANK and SUPPRES values in the distributed file are those used in Metathesaurus production. Users are free to change these values to suit their needs and preferences, then change the naming precedence and suppressibility (TS in MRCON) by using MetamorphoSys to create a customized Metathesaurus.

Col.	Description
RANK	Numeric order of precedence, higher value wins
SAB	Abbreviation for source vocabulary
TTY	Abbreviation for concept name type in source vocabulary
SUPPRESS	Flag indicating that this SAB and TTY will create a TS=s MRCON entry; see TS

Sample Records

0210|AIR|SY|N|
0209|ULT|PT|N|
0208|CPT|PT|N|

2.7.2.3.15 Ambiguous Term Identifiers (File = AMBIG.LUI)

In the instance that a Lexical Unique Identifier (LUI) is linked to multiple Concept Unique Identifiers (CUIs), there is one row in this table for each LUI-CUIs pair. This file identifies those lexical variant classes which have multiple meanings in the Metathesaurus.

In the Metathesaurus, the LUI links all strings within the English language that are identified as lexical variants of each other by the luinorm program found in the UMLS SPECIALIST Lexicon and Tools. LUIs are assigned irrespective of the meaning of each string. This table may be useful to system developers who wish to make use of the lexical programs in their applications to identify and disambiguate ambiguous terms.

Col.	Description
LUI	Lexical Unique Identifier
CUI	Concept Unique Identifier

Sample Records

L0000003|C0010504|
L0000003|C0917995|
L0000032|C0010206|

2.7.2.3.16 Ambiguous String Identifiers (File=AMBIG.SUI)

In the instance that a String Unique Identifier (SUI) is linked to multiple Concept Unique Identifiers (CUIs), there is one row in this table for each SUI-CUIs pair.

This file resides in the META directory. In the Metathesaurus, there is only one SUI for each unique string within each language, even if the string has multiple meanings. This table is only of interest to system developers who make use of the SUI in their applications or in local data files.

Col.	Description

| SUI | String Unique Identifier |
| CUI | Concept Unique Identifier |

Sample Records

S0063890|C0026667|
S0063890|C1135584|
S0229413|C0008802|

2.7.2.3.17 Metathesaurus Change Files

There are six files or relations that identify key differences between entries in the previous and the current edition of the Metathesaurus. Developers can use these special files to determine whether there have been changes that affect their applications.

The usefulness of individual files will depend on how data from the Metathesaurus have been linked or incorporated in a particular application.

Each relation or named table of data has a fixed number of columns and variable number of rows. A column is a sequence of all the values in a given data element. A row contains the values for two or more data elements for one entry. The values for the different data elements in the row are separated by vertical bars (|). Each row ends with a vertical bar and line termination.

2.7.2.3.17.1 Deleted Concepts (File=DELETED.CUI)

There is exactly one row in this table for each reviewed concept that was present in the previous Metathesaurus and is not present in the current Metathesaurus.

Cols.
CUI Concept Unique Identifier in the previous Metathesaurus
STR Preferred name of this concept in the previous Metathesaurus

2.7.2.3.17.2 Merged Concepts (File=MERGED.CUI)

There is exactly one row in this table for each released concept in the previous Metathesaurus (CUI1) that was merged into another released concept from the previous Metathesaurus (CUI2). When this merge occurs, the first CUI (CUI1) was retired; this table shows the CUI (CUI2) for the merged concept in this Metathesaurus.

Entries in this file represent concepts pairs that were considered to have different meanings in the previous edition, but which are now identified as synonyms

Cols.
CUI1 Concept Unique Identifier in the previous Metathesaurus
CUI2 Concept Unique Identifier in this Metathesaurus in format C#######

2.7.2.3.17.3 Deleted Terms (File=DELETED.LUI)

There is exactly one row in this table for each Lexical Unique Identifier (LUI) that appeared in the previous version of the Metathesaurus, but does not appear in this version.

Metathesaurus Lexical Unique Identifiers (LUIs) are assigned by the luinorm program, part of LVG program in the UMLS SPECIALIST Lexicon and Tools; see Section 4 in this manual.

These entries represent the cases where LUIs identified by the previous release's luinorm program, when used to identify lexical variants in the previous Metathesaurus, are no longer found with this release's luinorm on this release's Metathesaurus. This does not necessarily imply the deletion of a string or a concept from the Metathesaurus.

Cols.
LUI Lexical Unique Identifier in the previous Metathesaurus
STR Preferred Name of Term in the previous Metathesaurus

2.7.2.3.17.4 Merged Terms (File=MERGED.LUI)

There is exactly one row in this file for each case in which strings had different Lexical Unique Identifiers (LUIs) in the previous Metathesaurus yet share the same LUI in this Metathesaurus; a LUI present in the previous Metathesaurus is therefore absent from this Metathesaurus.

Metathesaurus Lexical Unique Identifiers (LUIs) are assigned by the luinorm program, part of the LVG program in the UMLS SPECIALIST Lexicon and Tools; see Sections 4 and 4.8 in this manual.

These entries represent the cases where separate lexical variants as identified by the previous release's luinorm program version are a single lexical variant as identified by this release's luinorm.

Cols.
LUI Lexical Unique Identifier in the previous Metathesaurus but not present in this Metathesaurus
LUI Lexical Unique Identifier into which it was merged in this Metathesaurus

2.7.2.3.17.5 Deleted Strings (File=DELETED.SUI)

There is exactly one row in this file for each string in each language that was present in a entry in the previous Metathesaurus and does not appear in this Metathesaurus.

Note that this does not necessarily imply the deletion of a term (LUI) or a concept (CUI) from the Metathesaurus. A string deleted in one language may still appear in the Metathesaurus in another language.

Cols.
SUI String Unique Identifier in previous Metathesaurus that is not present in this Metathesaurus
LAT Three character abbreviation of language of string that has been deleted.
STR Preferred name of term in previous Metathesaurus that is not present in this Metathesaurus.

2.7.2.3.17.6 Retired CUI Mapping (File=MRCUI)

There are one or more rows in this file for each Concept Unique Identifier (CUI) that existed in any prior release but is not present in the current release. The file includes mappings to current CUIs as synonymous or to one or more related current CUI where possible. If a synonymous mapping can not be found, other relationships between the CUIs can be created. These relationships can be Broader (RB), Narrower (RN), Other Related (RO), Deleted (DEL) or Removed from Subset (SUBX). Rows with the SUBX relationship are added to MRCUI by MetamorphoSys for each CUI that met the exclusion criteria and was consequently removed from the subset. Some CUIs may be mapped to more than one other CUI using these relationships.

CUIs may be retired when (1) two released concepts are found to be synonyms and so are merged, retiring one CUI; (2) when the concept no longer appears in any source vocabulary and is not 'rescued' by NLM; or (3) where the concept is an acknowledged error in a source vocabulary or determined to be a Metathesaurus production error.

See the META/CHANGE files, especially MERGED.CUI and DELETED.CUI, for the changes from the last release only, without mappings.

Col.	Description
CUI1	Retired CUI - was present in some prior release, but is currently missing
VER	The last release version in which CUI1 was a valid CUI
CREL	The relationship CUI2 has to CUI1, if present, or DEL if CUI2 is not present. Valid values currently are SY,DEL, RO, RN, RB
CUI2	The current CUI that CUI1 most closely maps to.
MAPIN	Is this map in current subset? Values of Y or N or null. MetamorphoSys generates the Y or N to indicate whether the CUI2 concept is or is not present in the subset. The null value is for rows where the CUI1 was not present to begin with (i.e., REL=DEL).

Sample Records:
C0079138|2001AA|DEL||Y|
C0079138|2001AA|RO|C0037440|Y|
C0079151|1993AA|DEL||N|
C0079158|1997AA|SY|C0009081||
C0079167|1997AA|SY|C0010042|N|

2.7.2.3.18 Word Index (File = MRXW.BAQ, MRXW.DAN, MRXW.DUT, MRXW.ENG, MRXW.FIN, MRXW.FRE, MRXW.GER, MRXW.HEB, MRXW.HUN, MRXW.ITA, MRXW.NOR, MRXW.POR, MRXW.RUS, MRXW.SPA, MRXW.SWE)

There is one row in these tables for each word found in each unique Metathesaurus string (ignoring upper-lower case). All Metathesaurus entries have entries in the word index. The entries are sorted in ASCII order.

Col.	Description
LAT	Abbreviation of language of the string in which the word appears
WD	Word in lowercase
CUI	Concept identifier
LUI	Term identifier
SUI	String identifier

Sample Records from MRXW.ENG

ENG|anaemia|C0002871|L0280031|S0352688|
ENG|anemia|C0002871|L0002871|S0013742|
ENG|anemias|C0002871|L0002871|S0013787|
ENG|blood|C0002871|L0376533|S0500659|

Sample Records from MRXW.FRE

FRE|ANEMIE|C0002871|L0162748|S0227229|

2.7.2.3.19 Normalized Word Index (File=MRXNW.ENG)

There is one row in this table for each normalized word found in each unique English-language Metathesaurus string. All English-language Metathesaurus entries have entries in the normalized word index. There are no normalized string indexes for other languages in this edition of the Metathesaurus.

Col.	Description
LAT	Abbreviation of language of the string in which the word appears (always ENG in this edition of the Metathesaurus)
NWD	Normalized word in lowercase (described in Section 2.6.2.1)
CUI	Concept identifier
LUI	Term identifier
SUI	String identifier

Sample Records

ENG|anaemia|C0002871|L0280031|S0352688|
ENG|anemia|C0002871|L0002871|S0013742|
ENG|anemia|C0002871|L0002871|S0013787|
ENG|blood|C0002871|L0376533|S0500659|

2.7.2.3.20 Normalized String Index (File=MRXNS.ENG)

There is one row in this table for each normalized string found in each unique English-language Metathesaurus string (ignoring upper-lower case). All English-language Metathesaurus entries have entries in the normalized string index. There are no normalized word indexes for other languages in this edition of the Metathesaurus.

Col.	Description
LAT	Abbreviation of language of the string (always ENG in this edition of the Metathesaurus)
NSTR	Normalized string in lowercase (described in Section 2.6.3.1)
CUI	Concept identifier
LUI	Term identifier
SUI	String identifier

Sample Records

ENG|anaemia|C0002871|L0280031|S0352688|
ENG|anaemia unspecified|C0002871|L0696700|S0803315|
ENG|anemia|C0002871|L0002871|S0013787|

2.8 Character Sets

Previous releases of the Metathesaurus used a simple 'least common denominator' character set known as '7-bit ASCII' or 'Basic Latin.' This standard contains no diacritics or special symbols, and is the default output encoding from MetamorphoSys.

Beginning with the 2004AA release, extended characters (including diacritical marks, ideographs, and scientific and other symbols) are also supported in Unicode, specifically in the the UTF-8 format of the Unicode 4.0 standard [1]. Unicode is the emerging international data encoding standard, and currently represents 96,382 different characters from the world's scripts and most languages.

Users may elect to output extended characters in MetamorphoSys. Further, users may choose to convert UTF-8 to other character sets using tools and online data tables available at http://www.unicode.org. The UMLS does not include character set conversion tools.

When extended characters appear in a source string, they are converted to UTF-8 in the Metathesaurus as necessary. For English sources, i.e., LAT =ENG, an equivalent 7-bit ASCII string is also created for the UMLS using the lvg program (see http://umlslex.nlm.nih.gov) to ensure that no information is lost when using the 7-bit ASCII character set.

2.8.1 7-bit ASCII or Basic Latin (used in Original Release Format)

This is the 'least common denominator' character set of 96 characters and symbols from the original ASCII standard. The UMLS 7-bit ASCII characters include those from 32 decimal to 127 decimal and these "C0 Controls" shown below:

0x09 CHARACTER TABULATION, horizontal tabulation. HT

0x0A LINE FEED, new line (NL), end of line (EOL)

0x0D CARRIAGE RETURN

This is the character set used in the original 'MR' format (Original Release Format) files, compatible with most computer systems.

2.8.2 Unicode and UTF-8 (optional extended characters; used in Rich Release Format, beginning with the UMLS 2004AA Release)

Only users who are familiar with Unicode and have fully compliant systems should enable the UTF-8 output of extended characters in MetamorphoSys.

Data from some sources in the Metathesaurus is now encoded in UTF-8 (UTF=Unicode Transformational Format), an 8-bit encoding suitable for processing in byte-oriented computer systems. It is a variable length encoding, so that a character can span one or more bytes. The initial byte order mark (BOM) character is not present in the UTF-8 encoded Metathesaurus files.

Note that UTF-8 is identical to the ASCII encoding for characters in the 7-bit ASCII range, so that 7-bit ASCII files are automatically a correct subset of UTF-8; this subset is the MetamorphoSys default.

The NLM receives data in a variety of character sets from source providers. Typically, files are encoded in variations of ASCII, including the ISO 8859 or the Windows codepage 1250 families of character sets. When not supplied in UTF-8, NLM will convert new or updated Sources to UTF-8, as resources permit.

The Metathesaurus has historically contained names for concepts from languages other than English (the different translations of the MeSH vocabulary, for example), but the characters used to represent these names were coerced to ASCII using a transliteration scheme. When users need correct extended characters and when non-European sources are added to the UMLS, the information loss becomes unacceptable and transliteration may be completely impossible.

From the 2004AA release forward:

1. Names of concepts in all vocabularies will be represented as supplied but converted to Unicode when necessary.
2. A 7-bit transliteration of strings in Western European languages that contain extended characters will be added, using the new lvg flow (see below). These strings will be identified in the MRCONSO.RRF file with the SAB and TTY columns set to **INSERT SABs* to select or exclude them.
3. There will be a relationship linking the extended character strings and their 7-bit translations.
4. Files will be in byte sort order (with data in UTF-8, standard UNIX sort works as expected). Note that the UMLS data are intended to be manipulated with software tools such as database systems, so the sort order of the files should not matter.

Table of Contents

Section 3
SEMANTIC NETWORK

3.0 Introduction

The Semantic Network consists of (1) a set of broad subject categories, or **Semantic Types**, that provide a consistent categorization of all concepts represented in the UMLS Metathesaurus, and (2) a set of useful and important relationships, or **Semantic Relations**, that exist between Semantic Types. This section of the documentation provides an overview of the Semantic Network, and describes the files of the Semantic Network. Sample records illustrate structure and content of these files.

3.1 Overview

The purpose of the Semantic Network is to provide a consistent categorization of all concepts represented in the UMLS Metathesaurus and to provide a set of useful relationships between these concepts. All information about specific concepts is found in the Metathesaurus. The Network provides information about the set of basic semantic types, or categories, which may be assigned to these concepts, and it defines the set of relationships that may hold between the semantic types. The Semantic Network contains 135 semantic types and 54 relationships. The Semantic Network serves as an authority for the semantic types that are assigned to concepts in the Metathesaurus. The Network defines these types, both with textual descriptions and by means of the information inherent in its hierarchies.

The semantic types are the nodes in the Network, and the relationships between them are the links. There are major groupings of semantic types for organisms, anatomical structures, biologic function, chemicals, events, physical objects, and concepts or ideas. The current scope of the UMLS semantic types is quite broad, allowing for the semantic categorization of a wide range of terminology in multiple domains.

The Metathesaurus consists of terms from its source vocabularies. The meaning of each term is defined by its source, explicitly by definition or annotation; by context (its place in a hierarchy); by synonyms and other stated relationships between terms; and by its usage in description, classification, or indexing. Each Metathesaurus concept is assigned at least one semantic type. In all cases, the most specific semantic type available in the hierarchy is assigned to the concept. For example, the concept "Macaca" receives the semantic type " Mammal" because there is not a more specific type "Primate" available in the Network. The level of granularity varies across the Network. This has important implications for interpreting the meaning (i.e., semantic type) that has been assigned to a Metathesaurus concept. For example, a sub-tree under the node "Physical Object" is "Manufactured Object". It has only two child nodes, "Medical Device" and "Research Device". It is clear that there are manufactured objects other than medical devices and research devices. Rather than proliferate the number of semantic types to encompass multiple additional subcategories for these objects, concepts that are neither medical devices nor research devices are simply assigned the more general semantic type "Manufactured Object".

Figure 1 illustrates a portion of the Network. The semantic type "Biologic Function" has two children, "Physiologic Function" and "Pathologic Function", and each of these in turn has several children and grandchildren. Each child in the hierarchy is linked to its parent by the "isa" link.

The primary link in the Network is the "isa" link. This establishes the hierarchy of types within the Network and is used for deciding on the most specific semantic type available for assignment to a Metathesaurus concept. In addition, a set of non-hierarchical relations between the types has been identified. These are grouped into five major categories, which are themselves relations: "physically related to", "spatially related to", "temporally related to", "functionally related to", and "conceptually related to".

Figure 2 illustrates a portion of the hierarchy for Network relationships. The "affects" relationship, one of several functional relationships, has six children, including "manages", "treats", and "prevents".

The relations are stated between high level semantic types in the Network whenever possible and are generally inherited via the "isa" link by all the children of those types. Thus, for example, the relation "process of" is stated to hold between the semantic types "Biologic Function" and "Organism". Therefore, it also holds between "Organ or Tissue Function" (which is a "Physiologic Function", which is, in turn, a "Biologic Function") and "Animal" (which is an "Organism"). The relations are

stated between semantic types and do not necessarily apply to all instances of concepts that have been assigned to those semantic types. That is, the relation may or may not hold between any particular pair of concepts. So, though the relation "evaluation of" holds between the semantic types "Sign" and "Organism Attribute", a particular sign or a particular attribute may not be linked by this relation. Thus, signs such as "overweight" and "fever" are evaluations of the organism attributes "body weight" and "body temperature", respectively. However, "overweight" is not an evaluation of "body temperature", and "fever" is not an evaluation of "body weight".

In some cases there will be a conflict between the placement of types in the Network and the link to be inherited. If so, the inheritance of the link is said to be blocked. For example, by inheritance, the type "Mental Process" would be "process of" "Plant". Since plants are not sentient beings, this link is explicitly blocked. In other cases the nature of the relation is such that it should not be inherited by the children of the types that it links. In that case, the relation is defined for the two semantic types it explicitly links, but blocked for all the children of those types. For example, "conceptual part of" links "Body System" and "Fully Formed Anatomical Structure", but it should not link "Body System" to all the children of "Fully Formed Anatomical Structure", such as "Cell" or "Tissue".

Several portions of the MeSH hierarchy have been labeled with child to parent semantic relationships. All of the anatomy, diseases, and psychiatry and psychology sections have been labeled, as well as a portion of the biological sciences section. The links that are expressed between MeSH terms are, with a few exceptions, reflected in the Semantic Network. That is, if two MeSH terms are linked by a certain relation, then that link is expressed in the Network as a link between the semantic types that have been assigned to those MeSH terms. For example, "Amniotic Fluid", which is a "Body Substance", is a child of "Embryo", which is an "Embryonic Structure". The labeled relationship between "Amniotic Fluid" and its parent "Embryo" is "surrounds". This is allowable, since the relation "Body Substance surrounds Embryonic Structure" is represented in the Network.

Figure 3 shows a portion of the Semantic Network, illustrating the relations, either hierarchical or associative, that exist between semantic types.

The UMLS Semantic Network is provided in two formats: a relational table format and a unit record format.

3.2 Semantic Network ASCII Relational Format

There are two basic tables, two ancillary tables, and two bookkeeping tables included in this format. The two basic tables contain exactly the same information as the unit record file, but the information is presented differently. One table contains definitional information about the semantic types and relations; the other contains information about the structure of the Network. Each semantic type and each relation has been assigned a four character unique identifier (UI). These are of the form "T001", "T002", etc. The ancillary tables are expansions of the table that contains the Network structure. They give the fully inherited set of links represented in the Network. The first table is expressed as triples of UI's. The second is expressed as triples of names. The two bookkeeping tables describe the relational files and their fields. Fields in all tables are separated by a "|". All tables are listed and described below:

Table	Description
SRDEF	Basic information about the Semantic Types and Relations.
SRSTR	Structure of the Network.
SRSTRE1	Fully inherited set of Relations (UI's).
SRSTRE2	Fully inherited set of Relations (names).
SRFIL	Description of each table.
SRFLD	Description of each field and the table(s) in which it is found.

Specific Descriptions of each Table:

Table: SRDEF

RT: Record Type (STY = Semantic Type or RL = Relation).

UI:	Unique Identifier of the Semantic Type or Relation.
STY/RL:	Name of the Semantic Type or Relation.
STN/RTN:	Tree Number of the Semantic Type or Relation.
DEF:	Definition of the Semantic Type or Relation.
EX:	Examples of Metathesaurus concepts with this Semantic Type (STY records only).
UN:	Usage note for Semantic Type assignment (STY records only).
NH:	The Semantic Type and its descendants allow the non-human flag (STY records only).
ABR:	Abbreviation of the Relation Name or Semantic Type
RIN:	Inverse of the Relation (RL records only).

Table: SRSTR

STY/RL:	Argument 1 (Name of a Semantic Type or Relation).
RL:	Relation ("isa" or the name of a non-hierarchical Relation).
STY/RL:	Argument 2 (Name of a Semantic Type or Relation); if this field is blank this means that the Semantic Type or Relation is one of the top nodes of the Network.
LS:	Link Status (D = Defined for the Arguments and its children; B = Blocked; DNI = Defined but Not Inherited by the children of the Arguments). N.B.: The relations expressed in this table are binary relations and the arguments are ordered pairs. The relations are stated only for the top-most node of the "isa" hierarchy of the Semantic Types to which they may apply.

Table: SRSTRE1 or SRSTRE2

UI/STY:	Argument 1 (UI or name of a Semantic Type).
UI/RL:	Relation (UI or name of a nonhierarchical Relation).
UI/STY:	Argument 2 (UI or name of a Semantic Type). N.B.: The relations expressed in this table are binary relations and the arguments are ordered pairs. All relations have been fully inherited in this table.

Table: SRFIL

FIL:	File Name.
DES:	Description of the file.
FMT:	Format of the file (fields in a comma-separated list).
CLS:	Number of columns in the file.
RWS:	Number of rows in the file.
BTS:	Number of bytes in the file.

Table: SRFLD

COL:	Field name.
DES:	Description of the field.
REF:	Cross-reference to the documentation.
FIL:	File name(s) in which the field is found.

Sample Relational Records

```
:::::::::::::::
SRDEF
:::::::::::::::
STY|T020|Acquired Abnormality|A1.2.2.2|An abnormal structure,
or one that is abnormal in size or location, found in or deriving
from a previously normal structure. Acquired abnormalities are
distinguished from diseases even though they may result in pathological
functioning (e.g., "hernias incarcerate").|Abscess of
prostate; Hemorrhoids; Hernia, Femoral; Varicose Veins|||||
STY|T052|Activity|B1|An operation or series of operations that
an organism or machine carries out or participates in.|Social
Planning; Expeditions; Information Distribution; Return Migration|Few
concepts will be assigned to this broad type. Wherever possible,
one of the more specific types from this hierarchy will be chosen.
 For concepts assigned to this type, the focus of interest is
on the activity. When the focus of interest is the individual
or group that is carrying out the activity, then a type from the
'Behavior' hierarchy will be chosen. In general, concepts will
not receive a type from both the 'Activity' and the 'Behavior'
hierarchies.||||
STY|T100|Age Group|A2.9.4|An individual or individuals classified
according to their age.|Adult; Infant, Premature; Adolescents;
Aged, 80 and over|||||
STY|T003|Alga|A1.1.1.1|A chiefly aquatic plant that contains chlorophyll,
but does not form embryos during development and lacks vascular
tissue.|Chlorella; Laminaria; Seaweed|||||
RL|T173|adjacent_to|R2.2|Close to, near or abutting another physical
unit with no other structure of the same kind intervening. This
includes adjoins, abuts, is contiguous to, is juxtaposed, and
is close to.||||AD|adjacent_to|
RL|T151|affects|R3.1|Produces a direct effect on. Implied
is the altering or influencing of an existing condition, state,
situation, or entity. This includes has a role in, alters, influences,
predisposes, catalyzes, stimulates, regulates, depresses, impedes,
enhances, contributes to, leads to, and modifies.||||AF|affected_by|
:::::::::::::::
SRSTR
:::::::::::::::
Acquired Abnormality|co-occurs_with|Injury or Poisoning|D|
Acquired Abnormality|isa|Anatomical Abnormality|D|
Acquired Abnormality|result_of|Behavior|D|
Activity|isa|Event|D|
Age Group|isa|Group|D|
Alga|isa|Plant|D|
:::::::::::::::
SRSTRE1
:::::::::::::::
T020|T186|T190|
T020|T186|T017|
T020|T186|T072|
T052|T186|T051|
T052|T165|T090|
T052|T165|T091|
T100|T186|T096|
T100|T186|T077|
```

```
T100|T186|T071|
T003|T186|T002|
T003|T186|T001|
T003|T186|T072|
:::::::::::::::
SRSTRE2
:::::::::::::::
Acquired Abnormality|isa|Anatomical Abnormality|
Acquired Abnormality|isa|Anatomical Structure|
Acquired Abnormality|isa|Physical Object|
Acquired Abnormality|isa|Entity|
Acquired Abnormality|affects|Alga|
Acquired Abnormality|affects|Amphibian|
Acquired Abnormality|affects|Animal|
Acquired Abnormality|affects|Bacterium|
Acquired Abnormality|affects|Bird|
Acquired Abnormality|affects|Cell Function|
Acquired Abnormality|affects|Fish|
Acquired Abnormality|affects|Fungus|
Acquired Abnormality|affects|Genetic Function|
Acquired Abnormality|affects|Human|
Acquired Abnormality|affects|Invertebrate|
Acquired Abnormality|affects|Mammal|
Acquired Abnormality|affects|Mental Process|
Acquired Abnormality|affects|Molecular Function|
Acquired Abnormality|affects|Organ or Tissue Function|
Acquired Abnormality|affects|Organism Function|
Acquired Abnormality|affects|Organism|
Acquired Abnormality|affects|Physiologic Function|
Acquired Abnormality|affects|Plant|
Acquired Abnormality|affects|Reptile|
Acquired Abnormality|affects|Rickettsia or Chlamydia|
Acquired Abnormality|affects|Vertebrate|
Acquired Abnormality|affects|Virus|
Activity|isa|Event|
Age Group|isa|Group|
Age Group|isa|Conceptual Entity|
Age Group|isa|Entity|
Alga|isa|Plant|
Alga|isa|Organism|
Alga|isa|Physical Object|
Alga|isa|Entity|
```

3.3 Semantic Network ASCII Unit Record Format

The file "SU" contains individual records for both semantic types and relations.

Each record begins with a unique identifier field (UI) which contains the four character UI. These are of the form "T001", "T002", etc. Each field in a record begins on a new line and may continue over several lines. Some fields are optional.

Semantic Type records contain the following fields:

Field	Description
UI:	Unique Identifier of the Semantic Type.
STY:	Name of the Semantic Type.

STN: Tree Number of the Semantic Type.

DEF: Definition of the Semantic Type.

EX: Examples of Metathesaurus concepts with this Semantic Type (optional field).

UN: Usage note for Semantic Type assignment (optional field).

NH: Semantic Type and its descendants allow the non-human flag (optional field).

HL: Hierarchical links of the Semantic Type to its parent({isa})and its children ({inverse_isa}). If there are no hierarchical links, then the value <none> is assigned.

Relation records contain the following fields:

Field	Description
UI:	Unique Identifier of the Relation.
RL:	Name of the Relation.
ABR:	Abbreviation of the Relation.
RIN:	Name of the inverse of the Relation.
RTN:	Tree Number of the Relation.
DEF:	Definition of the Relation.
INH:	"N" if the relation is not inherited (optional field).
HL:	Hierarchical links of the Relation to its parent ({isa}) and its children ({inverse_isa}). If there are no hierarchical links, then the value <none> is assigned.
STL:	Semantic Types linked by this Relation. N.B.: These are binary relations and the arguments are ordered pairs. The relations are stated only for the top-most node of the "isa" hierarchy of the Semantic Types to which they may apply. This field does not appear in the "isa" relation record since its values can be computed from the "HL" field. If there are no semantic types linked by this Relation, then the value <none> is assigned.
STLB:	Semantic Types linked by this Relation are blocked (optional field).

Sample Unit Records

```
::::::::::::::
SU
::::::::::::::
UI:  T020
STY: Acquired Abnormality
STN: A1.2.2.2
DEF: An abnormal structure, or one that is abnormal in size or location,
     found in or deriving from a previously normal structure. Acquired
     abnormalities are distinguished from diseases even though they may
     result in pathological functioning (e.g., "hernias incarcerate").
EX:  Abscess of prostate; Hemorrhoids; Hernia, Femoral; Varicose Veins
HL:  {isa} Anatomical Abnormality


UI:  T052
STY: Activity
STN: B1
DEF: An operation or series of operations that an organism or machine
     carries out or participates in.
EX:  Social Planning; Expeditions; Information Distribution; Return Migration
UN:  Few concepts will be assigned to this broad type. Wherever possible, one of the
     more specific types from this hierarchy will be chosen. For concepts assigned
```

to this type, the focus of interest is on the activity. When the focus of
interest is the individual or group that is carrying out the activity, then a
type from the 'Behavior' hierarchy will be chosen. In general, concepts will
not receive a type from both the 'Activity' and the 'Behavior' hierarchies.
HL: {isa} Event;
     {inverse_isa} Behavior;
     {inverse_isa} Daily or Recreational Activity;
     {inverse_isa} Occupational Activity;
     {inverse_isa} Machine Activity

UI: T100
STY: Age Group
STN: A2.9.4
DEF: An individual or individuals classified according to their age. EX: Adult;
     Infant, Premature; Adolescents; Aged, 80 and over
HL: {isa} Group

UI: T003
STY: Alga
STN: A1.1.1.1
DEF: A chiefly aquatic plant that contains chlorophyll, but does not form embryos
     during development and lacks vascular tissue.
EX: Chlorella; Laminaria; Seaweed
HL: {isa} Plant

UI: T173
RL: adjacent_to
ABR: AD
RIN: adjacent_to
RTN: R2.2
DEF: Close to, near or abutting another physical unit with no other structure of
     the same kind intervening. This includes adjoins, abuts, is contiguous to,
     is juxtaposed, and is close to.
HL: {isa} spatially_related_to
STL:
     [Body Location or Region|Body Location or Region];
     [Body Location or Region|Body Part, Organ, or Organ Component];
     [Body Location or Region|Body Space or Junction];
     [Body Part, Organ, or Organ Component|Body Part, Organ, or Organ Component];
     [Body Part, Organ, or Organ Component|Body Space or Junction];
     [Body Part, Organ, or Organ Component|Cell];
     [Body Part, Organ, or Organ Component|Tissue];
     [Body Space or Junction|Body Space or Junction];
     [Cell Component|Body Space or Junction];
     [Cell Component|Cell Component];
     [Cell|Cell];
     [Tissue|Body Space or Junction];
     [Tissue|Tissue]

UI: T151
RL: affects
ABR: AF
RIN: affected_by
RTN: R3.1
DEF: Produces a direct effect on. Implied here is the altering or influencing of
     an existing condition, state, situation, or entity. This includes has a role

```
        in, alters, influences, predisposes, catalyzes, stimulates, regulates,
        depresses, impedes, enhances, contributes to, leads to, and modifies.
HL:  {isa} functionally_related_to;
        {inverse_isa} manages;
        {inverse_isa} treats;
        {inverse_isa} disrupts;
        {inverse_isa} complicates;
        {inverse_isa} interacts_with;
        {inverse_isa} prevents
STL:
        [Natural Phenomenon or Process|Natural Phenomenon or Process];
        [Anatomical Abnormality|Physiologic Function];
        [Biologic Function|Organism];
        [Anatomical Abnormality|Organism];
        [Health Care Activity|Biologic Function];
        [Diagnostic Procedure|Patient or Disabled Group];
        [Therapeutic or Preventive Procedure|Patient or Disabled Group];
        [Chemical|Natural Phenomenon or Process];
        [Gene or Genome|Physiologic Function];
        [Cell Component|Physiologic Function];
        [Physiologic Function|Organism Attribute];
        [Food|Biologic Function];
        [Behavior|Behavior];
        [Behavior|Mental Process];
        [Mental Process|Behavior];
        [Mental or Behavioral Dysfunction|Behavior];
        [Research Activity|Mental Process];
        [Regulation or Law|Group];
        [Regulation or Law|Organization]
```

3.4 Hierarchies for Semantic Types and Relations In the Semantic Network

Current Semantic Types

Current relations in the Semantic Network

Semantic Network Web Site

Semantic Network terms and conditions for use

Table of Contents

Section 4
SPECIALIST LEXICON AND LEXICAL TOOLS

Introduction

The SPECIALIST lexicon has been developed to provide the lexical information needed for the SPECIALIST Natural Language Processing System (NLP). It is intended to be a general English lexicon that includes many biomedical terms. Coverage includes both commonly occurring English words and biomedical vocabulary. The lexicon entry for each word or term records the syntactic, morphological, and orthographic information needed by the SPECIALIST NLP System.

The [lexical tools](#) are designed to address the high degree of variability in natural language words and terms. Words often have several inflected forms which would properly be considered instances of the same word. The verb "treat", for example, has three inflectional variants: "treats " the third person singular present tense form, "treated" the past and past participle form, and "treating" the present participle form. Multi-word terms in the Metathesaurus and other controlled vocabularies may have word order variants in addition to their inflectional and alphabetic case variants. The lexical tools allow the user to abstract away from this sort of variation.

For an overview of the SPECIALIST lexicon, lexical variant programs, and lexical databases, see "Lexical Methods for Managing variation in Biomedical Terminologies", A.T. McCray, S. Srinivasan, A.C. Browne, in the Proceedings of the 18th Annual Symposium on Computer Applications in Medical Care, 1994, 235-239.

The SPECIALIST lexicon is distributed as one of the UMLS Knowledge Sources and as an open source resource along with the [the SPECIALIST NLP tools](#), subject to these [terms and conditions](#).

4.1 General Description

The lexicon consists of a set of lexical entries with one entry for each spelling or set of spelling variants in a particular part of speech. Lexical items may be "multi-word" terms made up of other words if the multi-word term is determined to be a lexical item by its presence as a term in general English or medical dictionaries, or in medical thesauri such as MeSH. Expansions of generally used acronyms and abbreviations are also allowed as multi-word terms.

The unit lexical record is a frame structure consisting of slots and fillers. Each lexical record has a base= slot whose filler indicates the base form, and optionally a set of spelling_variants= slots to indicate spelling variants. An "entry=" slot records the unique identifier (EUI) of the record. EUI numbers are seven digit numbers preceded by an "E". Each record has a cat= slot indicating part of speech. The lexical record is delimited by braces ({...}).

The unit lexical records for "anaesthetic" given below illustrate some of the features of the SPECIALIST lexical record:

{base=anesthetic spelling_variant=anaesthetic entry=E0330018 cat=noun variants=reg variants=uncount } {base=anesthetic spelling_variant=anaesthetic entry=E0330019 cat=adj variants=inv position=attrib(3) position=pred stative }

The base form "anesthetic" and its spelling variant "anaesthetic" appear in two lexical records, one an adjective entry, the other a noun entry. The variants= slot contains a code indicating the inflectional morphology of the entry; the filler reg in the noun entry indicates that the noun "anesthetic" is a count noun which undergoes regular English plural formation ("anaesthetics"); inv in the variants= slot of the adjective entry indicates that the adjective "anesthetic" does not form a comparative or superlative. The position= slot indicates that the adjective "anaesthetic" is attributive and appears after color adjectives in the normal adjective order. "pred" in the position slot of the adjective entry indicates that this adjective can appear in predicate position.

Lexical entries are not divided into senses. Therefore, an entry represents a spelling-category pairing regardless of semantics. The noun "act" has two senses both which show a capitalized and lower case spelling; an act of a play and an act of law. Since both senses share the same spellings and syntactic category, they are represented by a single lexical entry in the current lexicon. The unit record for "Act" is shown below.

{base=Act spelling_variant=act entry=E0000154 cat=noun variants=reg }

When different senses have different syntactic behavior, codes for each behavior are recorded in a single entry. For example, "beer" has two senses: the alcoholic beverage and the amount of a standard container of that beverage.

> A. Patients who drank beer recovered more slowly than patients who drank wine.
> B. Fifty-six patients reported drinking more than five beers a day.

The first sense illustrated in A. above is a mass (uncount) noun. The second sense illustrated in B. is a regular (count) noun. In cases like this the appropriate codes for both senses are included in the entry.

{base=beer
entry=E0012226 cat=noun variants=uncount variants=reg }

Two codes will also appear in cases where the lexical item is both count and uncount without a sense distinction. "Abdominal delivery" denotes the same procedure whether it appears as an uncount noun as in C. or a count noun as in D.

      C. Abdominal delivery is the procedure of choice in this situation.
      D. Abdominal deliveries are more common these days.

The unit lexical record for "abdominal delivery" includes both codes.

{base=abdominal delivery entry=E0006453 cat=noun variants=uncount variants=reg }

Other syntactic codes such as complement codes for verbs, adjectives and nouns are similarly grouped without regard to sense.

4.2 The Scope of the Lexicon

Words are selected for lexical coding from a variety of sources. Approximately 20,000 words from the UMLS Test Collection of MEDLINE abstracts together with words which appear both in the UMLS Metathesaurus and Dorland's Illustrated Medical Dictionary form the core of the words entered. In addition, an effort has been made to include words from the general English vocabulary. The 10,000 most frequent words listed in The American Heritage Word Frequency Book and the list of 2,000 words used in definitions in Longman's Dictionary of Contemporary English have also been coded. Since the majority of the words selected for coding are nouns, an effort has been made to include verbs and adjectives by identifying verbs in current MEDLINE citation records, by using the Computer Usable Oxford Advanced Learner's Dictionary, and by identifying potential adjectives from Dorland's Illustrated Medical Dictionary using heuristics developed by McCray and Srinivasan (1990).

A variety of reference sources are used in coding lexical records. Coding is based on actual usage in the UMLS Test Collection and MEDLINE, dictionaries of general English, primarily learner's dictionaries which record the kind of syntactic information needed for NLP, and medical dictionaries. Longman's Dictionary of Contemporary English, Dorland's Illustrated Medical Dictionary, Collins COBUILD Dictionary, The Oxford Advanced Learner's Dictionary, and Webster's Medical Desk Dictionary were used.

The SPECIALIST lexicon also exists in relational format generated from the unit records. The full SPECIALIST lexicon technical report entitled "The SPECIALIST Lexicon", found in the file techrpt.pdf, fully describes the unit record format. The remainder of the present section describes the relational form of the lexicon. Section 3 describes the Data elements that make up the relational tables and Section 4 describes the tables.

4.3 Lexicon Data Elements
Each of the elements below are represented as fields (columns) in the relational format.

4.3.1 String Properties

These data elements refer to properties of the strings generated by the entries.

4.3.1.1 STR - String

A Lexical entry generates a variety of forms (strings) including all the inflectional forms (the citation form, as well) of each spelling variant. Case, punctuation and spaces are considered significant.

4.3.1.2 AGR - Agreement/Inflection Code

This element encodes agreement and inflection information.

Agreement between nouns and verbs and between determiners and nouns involves person and number. Person and Number are indicated by the following codes.

Code	Person	Number
second	Second	Singular & Plural
third	Third	Singular & Plural
fst_sing	First	Singular
fst_plur	First	Plural
thr_sing	Third	Singular
thr_plur	Third	Plural

For Nouns, the agreement/inflection code indicates countability, person and number. Person and number are indicated by the person/number codes given above which are parenthesized after the countability code. Nouns can be either count or uncount.

For Pronouns, the agreement/inflection indicates person and number using the codes given above.

For verbs, including auxiliaries and modals, the agreement/inflection code indicates tense, person and number. Persons and numbers are indicated by the same person/number codes given above. These codes are parenthesized after the tense. No person number codes are given for non-finite tenses. "pres(thr_sing)" indicates third person singular present tense and "pres (fst_sing,fst_plur,thr_plur,second)" indicates present tense for all persons and numbers other than third singular. Negative forms of auxiliaries (didn't) and modals (can't) have "negative" after a colon at the end of the agreement/inflection code.

Code	Tense
past	Past Tense
pres	Present Tense
past_part	Past Participle
pres_part	Present Participle
infinitive	Infinitive

Determiners agree with nouns in terms of countability and number. The agreement/inflection codes for determiners are "free", "plur", "sing" and "uncount". "free" indicates that the determiner places no restrictions on its noun. Determiners marked "plur" allow plural nouns, those marked "sing" allow singular nouns and those marked "uncount" allow uncountnouns.

4.3.1.3 CAS - Case

See Section 4.3.1 of "The SPECIALIST Lexicon".

Pronouns in English may be in one of two cases, subjective (nominative) or objective (accusative). This field contains "subj", "obj" or both separated by a comma to indicate the case of the pronoun.

4.3.1.4 GND - Gender

This field indicates the gender of pronouns.

Pronouns may be marked pers or neut to indicate whether they refer to people or non-people respectively. Pronouns marked pers may be masculine (masc) or feminine (fem) referring to male or female people respectively. See Section 14.2 of "The SPECIALIST Lexicon". There are four codes possible in this field:

Code	Gender
pers	person
neut	neuter
pers(masc)	person masculine
pers(fem)	person feminine

Notice that pers as used here does not correspond to the traditional term "personal pronoun". For example "it" and "they" are traditionally called personal pronouns since they both participate in the person/number paradigm. A pronoun like "none" is not traditionally called a personal pronoun.

4.3.2 Entry Properties 4.3.2.1 EUI - Unique Identifier Number for Lexical Entries

The EUI identifies a lexical entry. Information about a set of spelling variants in a particular part of speech is represented as an entry in the unit record. A particular string may be assigned several EUI numbers as it may occur in several parts of speech.

4.3.2.2 CIT - Citation Form

This field records the citation form of strings in the agreement/inflection table (lragr). The citation form is the singular for nouns, infinitive for verb and positive for adjectives and adverbs. The base form and the spelling variants if any are the citation forms of each of their respective inflections. This form is sometimes referred to as the un-inflected form.

4.3.2.3 BAS - Base Form

This field records the base form of a lexical entry. The base form is the citation form of one of a set of spelling variants chosen to represent the whole set. It might be thought of as the name of a lexical entry. The base form is the filler of the base= slot.

4.3.2.4 SCA - Syntactic Category

The syntactic category (part of speech) of the lexical entry. This field may be filled by one of the following. See Section 3 of "The SPECIALIST Lexicon".

Code	Category
noun	nouns
adj	adjectives
adv	adverbs
pron	pronouns
verb	verbs

det	determiners
prep	prepositions
conj	conjunctions
aux	auxiliaries
modal	modals
compl	complementizers

4.3.2.5 PER - Periphrastic

The code "periph" in this field indicates that an adjective or adverb is periphrastic. An adjective is periphrastic if it can form its comparative with "more" and its superlative with "most". See Section 4.3.5 of "The SPECIALIST Lexicon" for discussion.

4.3.2.6 COM - Complements

These are complement codes. See Sections 5.1, 5.2, 5.4 and 5.5 in "The SPECIALIST Lexicon" for a description of SPECIALIST complement codes.

4.3.2.7 TYP - Inflectional Type

The inflectional type(s) of an entry indicate the ways in which it's forms may be inflected, or in the case of determiners the inflection of the heads they may determine. These codes are used to generate the variant strings (STR) found in other tables.

For nouns the following types may appear:

Code	Pluralization Pattern	See "The SPECIALIST Lexicon" Section
reg	regular	4.5.2
glreg	Greco-Latin regular	4.5.3
metareg	metalinguistic regular	4.5.4
irreg( )	irregular	4.5.5
sing	fixed singular	4.5.6
plur	fixed plural	4.5.7
inv	invariant	4.5.8
group(irreg( ))	group irregular	4.5.9
group(reg)	group regular	4.5.9
uncount	uncountable	4.5.10
groupuncount	group uncount	4.5.11

For verbs the following types may appear:

Code	Inflection Type	See "The SPECIALIST Lexicon" Section
reg	regular	4.1.1

regd	regular doubling	4.1.2
irreg( )	irregular	4.1.3

For pronouns the following types may appear:

Code	Inflection Type
fst_plur	first person plural
fst_sing	first person singular
sec_plur	second person plural
sec_sing	second person singular
second	second person
third	third person
thr_plur	third person plural
thr_sing	third person singular

See Section 14.1 of "The SPECIALIST Lexicon".

For adjectives and adverbs the following types can appear:

Code	Inflectional Type	See "The SPECIALIST Lexicon" Section
reg	regular	4.3.1 and 4.4.1
regd	regular doubling	4.3.2
inv	invariant	4.3.4 and 4.4.3
inv;periph	periphrastic	4.3.5 and 4.4.4
irreg( )	irregular	4.3.3 and 4.4.2

For determiners the infection type indicates the inflection of the noun heads they may determine. The following types may appear:

Code	Inflectional Type	See "The SPECIALIST Lexicon" Section
sing	singular	4.7.1
plur	plural	4.7.2
uncount	uncount	4.7.3
singuncount	singular uncount	4.7.4
pluruncount	plural uncount	4.7.5
free	free	4.7.6

4.3.2.8 POS - Possession

English pronouns may be possessive, or possessive nominal. The codes poss, possnom or both (comma separated) may appear in this field.

See Section 14.3.2 of "The SPECIALIST Lexicon".

4.3.2.9 QNT - Quantification

This field indicates the quantification properties inherent in certain pronouns. The four codes possible in this field are:

Code	Properties
univ	universal quantification
indef(nonassert)	non-assertive indefinite
indef(neg)	negative indefinite
indef(assert)	assertive indefinite

See Section 14.3.4 in "The SPECIALIST Lexicon" for discussion of quantification in pronouns.

4.3.2.10 FEA - Features

This field represents various features of terms in various categories. The possible features are:

Feature	Section of "The SPECIALIST Lexicon"
reflexive	14.3.3
negative	14.3.4
demonstrative	14.3.5
interrogative	12.1
proper	8.
negative	13.1
broad_negative	13.2
stative	10.

4.3.2.11 PSN - Position for Adjectives

Adjectives are marked in the SPECIALIST lexicon with position codes showing whether they are attributive postmodifying or predicative. If attributive, the code indicates where they appear in the pre-nominal sequence of adjectives. An additional attributive code, attribc, is used to indicate adjectives which can take complements in attributive position. One or more of the following codes can appear:

Code	Position	See "The SPECIALIST Lexicon" Section
attrib(1)	attributive (1st position)	9.1.1.1
attrib(2)	attributive (2nd position)	9.1.1.2

attrib(3)	attributive (3rd position)	9.1.1.3
attribc	attributive with complement	9.1.2
post	post modifying	9.2
pred	predicative	9.3

4.3.2.12 MOD - Modification Type for Adverbs

Adverbs are marked in the SPECIALIST lexicon to indicate their modification type. The possible values of this field are:

Code	See "The SPECIALIST Lexicon" Section
intensifier	11.2
particle	11.1
sentence_modifier; TYPE	11.3
verb_modifier; TYPE	11.4

TYPE is one of locative, temporal or manner. See Section 11.5 in "The SPECIALIST Lexicon".

4.3.2.13 GEN - Generic Name for a Trademark

The GEN field represents a generic or public name for the thing referred to by the trademark. The trademark "Alphalin" has the generic term "vitamin A".

4.3.3 Entry Relations 4.3.3.1 ABR - Acronym or Abbreviation

This field indicates whether a term listed in the acronym-abbreviation table (lrabr) is an acronym or abbreviation. It contains either:

"abbreviation_of" or "acronym_of".

4.3.3.2 SPV - Spelling Variant

A base form in the SPECIALIST lexicon may have one or more spelling variants, subject to the same inflectional pattern. This field contains the citation form of a particular spelling variant. See Section (2) of "The SPECIALIST Lexicon".

4.3.4 Data Description

The data elements describe the relational table files or provide index entries into the lexicon.

4.3.4.1 WRD - Word

Each string is broken into "words" and indexed in lrwd. Words are strings of alpha-numeric characters more than one character long, separated by space or punctuation.

4.3.4.2 DES - Description

A short definition of a file or field. This is free text.

4.3.4.3 FMT - Format

An ordered comma separated list of field names appearing in a file.

4.3.4.4 RWS - Number of Rows

The number of Rows (lines or records) in a file.

4.3.4.5 FIL - File Name(s)

One or more file names denoting the files containing relational tables.

4.3.4.6 BTS - Size in Bytes

The size of a file in bytes (characters).

4.3.4.7 CLS - Number of Columns

The number of columns (fields) in a record (or row) of a table. The same number as the number of lines in the file.

4.3.4.8 COL - Three Letter Field Name

A three letter identifier for a field.

4.3.4.9 REF - Cross Reference to Document

A cross reference to a section of this document.

4.4 Lexicon Relational Tables 4.4.1 Introduction

In this format the data in each lexical entry is represented in ten different "relations" or "tables" each in a file.

The Lexicon relational format is not fully normalized. By design, there is duplication of data among different relations and within certain relations. Developers will need to make their own decisions about the extent to which this redundancy should be retained, reduced, or increased for their specific applications.

4.4.2 General Description of the Relational Format

As in the Metathesaurus ASCII relational format, each relation or table of data values has by definition a fixed number of columns; the number of rows depends on the content of a particular version of the Lexicon. A column is a sequence of all the values in a given data element or logical sub-element. In general, columns for longer variable length data elements will appear to the right of columns for shorter and/or fixed length data elements. A row contains the values for one or more data elements or logical sub-elements for one Lexicon entry or string. Depending on the nature of the data elements involved, each Lexicon entry or string may have one or more rows in a given file. The values for the different data elements or logical sub-elements represented in the row are separated by vertical bars (|). If an optional element is blank, the vertical bars are still used to maintain the correct positioning of the subsequent elements. Each row is terminated by a vertical bar and a carriage return followed by a line feed. (|<CR><LF>).

4.4.3 Summary of the Contents of Each of the Relational Files

In the following descriptions, the numbers in parentheses beside each element refer to the section of this document that describes the element's contents.

4.4.3.1 - Agreement and Inflection (File = lragr)

Rows of the agreement table have six fields. There is a row in lragr for each inflected form of each spelling variant. This table links those forms to their citation forms and base forms. It provides information about agreement between subjects (nouns and pronouns) and verbs and between determiners and nouns.

EUI   The Entry Unique ID Number (3.2.1)

STR   String (3.1.1)

SCA   Syntactic Category (3.2.4)

AGR   Agreement/Inflection Code (3.1.2)

CIT   Citation Form (3.2.2)

BAS   Base Form (3.2.3)

4.4.3.2 - Inflection Type (File = lrtyp)

The lrtyp table has one or more rows for each lexical entry, indicating the inflectional pattern(s) to which it belongs.

EUI   The Entry Unique ID Number (3.2.1)

BAS   The Base Form (3.2.3)

SCA   Syntactic Category (3.2.4)

TYP   Inflectional Type (3.2.7)

4.4.3.3 - Complementation (File = lrcmp)

In lrcmp there is one line for each complement code for each entry.

EUI   The Entry Unique ID Number (3.2.1)

BAS   The Base Form (3.2.3)

SCA   Syntactic Category (3.2.4)

COM   Complement Code. (3.2.6)

4.4.3.4 - Pronouns (File = lrprn)

lrprn has one or more rows for each pronoun entry in the lexicon. Each row has nine columns.

EUI   The Entry Unique ID Number (3.2.1)

BAS   The Base Form (3.2.3)

AGR   Agreement/Inflection Code (3.1.2)

See Section 14.1 in "The SPECIALIST Lexicon".

The agreement/inflection field in lrprn indicates person and number for anaphoric reference, AGR in lragr indicates person for agreement. These differ in the case of possessive nominal pronouns. The possessive nominal "mine" is "third" for purposes of subject verb agreement and "fst_sing" in its anaphoric reference.

GND   Gender (3.1.6)

CAS   Case (3.1.3)

POS   Possession ([3.2.8](#))

QNT   Quantification ([3.2.9](#))

FEA   Other Features (for pronouns) ([3.2.10](#))

4.4.3.5 Modifiers (file = lrmod)

The modifier table includes position information for adjectives and modification type information for adverbs, and a variety of features.

EUI   The Entry Unique ID Number ([3.2.1](#))

BAS   The Base Form ([3.2.3](#))

SCA   Syntactic Category ([3.2.4](#))

All the entries represented in this table have the category "adj" or "adv" indicating adjectives or adverbs respectively.

PSN/MOD

The fourth field of lrmod may be one of the following depending on whether the term is an adjective or adverb.

PSN   Position ([3.2.11](#)) - for adjectives

MOD   Modification Types ([3.2.12](#)) - for adverbs

FEA   Features ([3.2.10](#))

4.4.3.6 - Properties (file = lrprp)

lrprp indicates properties of terms in various categories.

EUI   The Entry Unique ID Number ([3.2.1](#))

BAS   The Base Form ([3.2.3](#))

SCA   Syntactic Category ([3.2.4](#))

STR   String ([3.1.1](#))

STR is only indicated in lrprp when a feature applies to a single string out of those generated by the entry, as in the negative contractions.

FEA   Features (4.3.2.10)

4.4.3.7 - Abbreviations and Acronyms (file = lrabr)

This file links acronyms and abbreviations to their expansions.

EUI   The Entry Unique ID Number ([3.2.1](#))

This field contains the EUI of the acronym or abbreviation.

BAS   The Base Form ([3.2.3](#))

This field contains the Base form of the acronym or abbreviation.

ABR   Acronym or Abbreviation ([3.3.1](#))

BAS   The Base Form ([3.2.3](#))

This field contains the Base form of the expansion of the acronym or abbreviation.

EUI   The Entry Unique ID Number ([3.2.1](#))

This field contains the EUI of the expansion of the abbreviation or acronym.

4.4.3.8 - Spelling Variants (file = lrspl)

EUI   The Entry Unique ID Number ([3.2.1](#))

SPV   Spelling Variant ([3.3.2](#))

BAS   The Base Form ([3.2.3](#))

4.4.3.9 - Nominalizations (file = lrnom)

This field contains the EUI of the nominalization.

BAS   The Base Form ([3.2.3](#))

This field contains the base form of the nominalization.

SCA   Syntactic Category ([3.2.4](#))

This field contains the category of the nominalization (noun).

EUI   The Entry Unique ID Number ([3.2.1](#))

This field contains the EUI of a verb or adjective of which the noun is a nominalization.

BAS   The Base Form ([3.2.3](#))

This field contains the base form of the verb or adjective of which the noun is a nominalization.

SCA   Syntactic Category ([3.2.4](#))

This field contains the syntactic category (adj or verb) of the adjective or verb.

4.4.3.10 - Trademarks (file = lrtrm)

EUI   The Entry Unique ID Number ([3.2.1](#))

BAS   Base ([3.2.3](#))

GEN   Generic Term ([3.2.13](#))

The appearance of a form in the lrtrm table indicates that it is a trademark. It may or may not have a generic term associated with it.

4.4.3.11 - Files (file = lrfil)

The lrfil table describes each file in the ASCII relational form of the lexicon.

FIL   File Name(s) ([3.4.5](#))

DES   Description ([3.4.2](#))

FMT   Format ([3.4.3](#))

CLS   Number of Columns ([3.4.7](#))

RWS   Number of Rows ([3.4.4](#))

BTS   Size in Bytes ([3.4.6](#))

4.4.3.12 - Word Index. (file = lrwrd)

WRD   Word ([3.4.1](#))

EUI   The Entry Unique ID Number ([3.2.1](#))

4.4.3.13 - Fields (file = lrfld)

COL   Three Letter Field Name ([3.4.8](#))

DES   Description ([3.4.2](#))

REF   Cross Reference to Document ([3.4.9](#))

FIL   File Name(s) ([3.4.5](#))

4.5 The SPECIALIST Lexicon Unit Record

The unit lexical record is a frame structure consisting of slots and fillers. Each lexical record has a base= slot whose filler indicates the base form, and optionally a set of spelling_variants= slots to indicate spelling variants. Lexical entries are delimited by entry= slots filled by the EUI number of the entry. EUI numbers are seven digit numbers preceded by an "E". Each entry has a cat= slot indicating part of speech. The lexical record is delimited by braces ({...}).

The unit lexical records for "anaesthetic" given below illustrate some of the features of a SPECIALIST unit lexical record:

{base=anesthetic spelling_variant=anaesthetic entry=E0330018 cat=noun variants=reg variants=uncount } {base=anesthetic spelling_variant=anaesthetic entry=E0330019 cat=adj variants=inv position=attrib(3) position=pred stative }

The base form "anesthetic" and its spelling variant "anaesthetic" appear in two lexical records containing a noun and a verb entry. The variants= slot contains a code indicating the inflectional morphology of the entry; the filler reg in the noun entry indicates that the noun "anaesthetic" is a count noun which undergoes regular English plural formation ("anaesthetics"); inv in the variants= slot of the adjective entry indicates that the adjective "anesthetic" does not form a comparative or superlative. The position= slot indicates that the adjective "anaesthetic" is attributive and appears after color adjectives in the normal adjective order.

The SPECIALIST Technical report "The SPECIALIST Lexicon" gives a full description of the lexicon in unit format.

4.6 Lexical Databases Introduction

The lexical databases contain lexical information that we have found to be useful for Natural Language Processing. They are not finished products but are under continuous development.

4.6.1 Semantically Related Terms SM.DB

This database (SM.DB) contains pairs of semantically related terms. Each row of the database has the following form.

     TERM1|SCA1|TERM2|SCA2

Such a row indicates that TERM1 in syntactic category SCA1 is semantically related to TERM2 in syntactic category SCA2. Both terms are given in base form.

Examples:

alar|adj|wing|noun ocular|adj|eye|noun auditory area|noun|auditory cortex|noun vomitive|noun|emetic|noun vomitive|adj|emetic|adj iridescent virus|noun|iridovirus|noun typhloteritis|noun|cecitis|noun

4.6.2 Derivationally Related Terms: DM.DB

This database (DM.DB) contains pairs of terms related by derivational morphology. Each row of the database has the same form as sm.db. Both terms are given in base form.

TERM1|SCA1|TERM2|SCA2

Examples:

abashment|noun|abash|verb
adenohypophyseal|adj|adenohypophysis|noun
amenorrheic|adj|amenorrhea|noun
arithmetician|noun|arithmetic|noun
convert|verb|conversible|adj
immobilize|verb|immobility|noun

DM.DB is derived from the morphological fact files (dm.fct, etc.) used in LVG.

4.6.3 Spelling Variants: SP.DB

The Spelling Variant database (SP.DB) contains pairs of terms that are spelling variants of each other. The format of each row is the same as the format of dm.db and sm.db. SCA1 and SCA2 are always the same in SP.DB.

TERM1|SCA1|TERM2|SCA2

Examples:

accouter|verb|accoutre|verb
accurst|adj|accursed|adj
acidaemic|adj|acidemic|adj
aesthetics|noun|esthetics|noun
dairy farmer|noun|dairy-farmer|noun

SP.DB is derived from the SPECIALIST lexicon.

4.6.4 Neo-classical Combining Forms NC.DB

This database (NC.DB) contains morphemes that are used to form neo-classical compounds. Each row of the database has the following form.

MORPHEME|MEANING|TYPE

Morphemes may have optional connecting vowels indicated in parentheses. The types are: prefix, root, and terminal.

Examples:

```
abdomin(o)|abdomen|root
ab|away from|prefix
acou(o)|hearing|root
cardi(o)|heart|root
cele|swelling|terminal
desis|binding|terminal
de|negate|prefix
```

Our analysis of combining forms divides them into roots and terminals, which are distinguished from prefixes and suffixes. A neo-classical compound can consist of any number of roots ending in a terminal or suffix. Prefixes normally must precede roots and cannot attach directly to terminals. Users interested in suffixation rules and facts should consult the dm.rul and dm. fct files included with LVG.

For further discussion see McCray et. al., 1988, "The Semantic Structure of Neo-Classical Compounds", In the Proceedings of the Twelfth Annual Symposium on Computer Applications in Medical Care, Washington DC.

4.7 Sample Records

```
:::::::::::::
lragr.sam
:::::::::::::
E0007127|acute|adj|positive;periph|acute|acute|
E0014875|cans|noun|count(thr_plur)|can|can|
E0014875|can|noun|count(thr_sing)|can|can|
E0014876|canned|verb|past_part|can|can|
E0014876|canned|verb|past|can|can|
E0014876|canning|verb|pres_part|can|can|
E0014876|cans|verb|pres(thr_sing)|can|can|
E0014876|can|verb|infinitive|can|can|
E0014876|can|verb|pres(fst_sing,fst_plur,thr_plur,second)|can|can|
E0014877|can't|modal|pres:negative|can|can|
E0014877|cannot|modal|pres:negative|can|can|
E0014877|can|modal|pres|can|can|
E0014877|couldn't|modal|past:negative|can|can|
E0014877|could|modal|past|can|can|
E0014937|canine teeth|noun|count(thr_plur)|canine tooth|canine tooth|
E0014937|canine tooth|noun|count(thr_sing)|canine tooth|canine tooth|
E0017902|colors|noun|count(thr_plur)|color|color|
E0017902|color|noun|count(thr_sing)|color|color|
E0017902|color|noun|uncount(thr_sing)|color|color|
E0017903|colored|verb|past_part|color|color|
E0017903|colored|verb|past|color|color|
E0017903|coloring|verb|pres_part|color|color|
E0017903|colors|verb|pres(thr_sing)|color|color|
E0017903|color|verb|infinitive|color|color|
E0017903|color|verb|pres(fst_sing,fst_plur,thr_plur,second)|color|color|
E0051632|quickly|adv|positive;periph|quickly|quickly|
E0055585|she|pron|thr_sing|she|she|

:::::::::::::
lrcmp.sam
:::::::::::::
E0014876|can|verb|tran=np|
E0017903|color|verb|cplxtran=np,adj|
E0017903|color|verb|cplxtran=np,np|
E0017903|color|verb|intran;part(in)|
```

E0017903|color|verb|intran;part(up)|
E0017903|color|verb|intran|
E0017903|color|verb|tran=np;part(in)|
E0017903|color|verb|tran=np|

::::::::::::::
lrmod.sam
::::::::::::::
E0007127|acute|adj|attrib(1),attrib(3),pred|stative|
E0051632|quickly|adv|verb_modifier;manner||

::::::::::::::
lrnom.sam
::::::::::::::
E0007121|acuity|noun|E0007127|acute|adj|
E0021126|deduction|noun|E0021123|deduce|verb|
E0021126|deduction|noun|E0021124|deduct|verb|
E0061851|transportation|noun|E0061850|transport|verb|

::::::::::::::
lrprn.sam
::::::::::::::
E0030918|he|thr_sing|pers(masc)|subj|||||
E0036100|it|thr_sing|neut|subj,obj|||||
E0055585|she|thr_sing|pers(fem)|subj|||||

::::::::::::::
lrprp.sam
::::::::::::::
E0007127|acute|acute|adj|stative|
E0004825|Parkinson|Parkinson|noun|proper|
E0014877|can|can't|modal|negative|
E0014877|can|can't|modal|negative|
E0014877|can|couldn't|modal|negative|

::::::::::::::
lrspl.sam
::::::::::::::
E0017902|colour|color|
E0017903|colour|color|
E0008769|anesthetic|anaesthetic|
E0008770|anesthetic|anaesthetic|

::::::::::::::
lrtrm.sam
::::::::::::::
E0412633|Actinex|meso-nordihydroguaiaretic acid|
E0415286|Antivert||
E0414928|thioplex|thiotepa|
E0415019|theo-hexanicit||

::::::::::::::
lrtyp.sam
::::::::::::::
E0007127|acute|adj|inv;periph|
E0014875|can|noun|reg|
E0014876|can|verb|regd|

E0014937|canine tooth|noun|irreg(canine teeth)|
E0017902|color|noun|reg|
E0017902|color|noun|uncount|
E0017903|color|verb|reg|
E0051632|quickly|adv|inv;periph|

:::::::::::::
lrwd.sam
:::::::::::::
acute|E0001203
acute|E0007127
acute|E0007130
acute|E0007131
acute|E0007132
acute|E0007133
acute|E0007134
acute|E0007135
acute|E0007136
acute|E0007137
acute|E0007138
acute|E0007139
acute|E0007140
acute|E0007141
acute|E0007142
acute|E0007143
acute|E0007144
acute|E0007145
acute|E0007146
acute|E0007147
acute|E0007148
acute|E0007149
acute|E0007150
acute|E0007151
acute|E0007152
acute|E0007153
acute|E0007154
acute|E0007155
acute|E0007156
acute|E0007157
acute|E0007158
acute|E0007159
acute|E0007160
acute|E0007161
acute|E0007162
acute|E0007163
acute|E0007164
acute|E0007165
acute|E0007166
acute|E0007167
acute|E0007168
acute|E0007169
acute|E0007170
acute|E0007171
acute|E0007172
acute|E0007173
acute|E0007174

acute|E0007175
acute|E0007176
acute|E0007177
acute|E0007178
acute|E0007179
acute|E0007180
acute|E0007181
acute|E0007182
acute|E0007183
acute|E0007184
acute|E0007185
acute|E0007186
acute|E0007187
acute|E0007188
acute|E0007189
acute|E0007190
acute|E0007191
acute|E0007192
acute|E0007193
acute|E0007194
acute|E0007195
acute|E0007196
acute|E0007197
acute|E0007198
acute|E0007199
acute|E0007200
acute|E0007201
acute|E0007202
acute|E0007203
acute|E0007204
acute|E0007205
acute|E0007206
acute|E0007207
acute|E0007208
acute|E0007209
acute|E0007210
acute|E0007211
acute|E0007212
acute|E0007213
acute|E0007214
acute|E0007215
acute|E0016430
acute|E0018044
acute|E0019256
acute|E0200089
acute|E0200090
acute|E0203254
acute|E0208423
acute|E0208433
acute|E0208452
acute|E0208475
acute|E0208494
acute|E0210443
acute|E0210574
acute|E0210575
acute|E0210576

```
acute|E0210642
acute|E0214476
acute|E0216615
acute|E0216616
acute|E0217176
acute|E0217376
acute|E0217551
acute|E0217756
acute|E0313307
acute|E0314926
acute|E0319558
acute|E0321232
acute|E0321304
acute|E0322005
acute|E0332592
acute|E0409630
acute|E0418090
acute|E0418484
acute|E0418485
acute|E0418705
acute|E0420121
acute|E0422597
acute|E0422634
acute|E0422824
acute|E0422825
can|E0014875
can|E0014876
can|E0014877
can|E0014875
can|E0014876
can|E0014877
```

4.8 The SPECIALIST Lexical Tools

The SPECIALIST lexical tools package consists of three primary programs -- a normalizer, a word index generator, and a lexical variant generator, together with a set of ancillary programs for normalization. This package is implemented in Java.

The SPECIALIST Lexical Tools and the SPECIALIST lexicon are distributed as one of the UMLS Knowledge Sources and along with the SPECIALIST NLP Tools as open source resources subject to these terms and conditions.

Updates and bug fixes to this versions may be found at the url: http://umlslex.nlm.nih.gov/lvg/current/.

The distributions come with install programs (for Solaris, Linux, and Window) and a ReadMe.txt file describing how to install and configure the lexical programs and providing a brief description of each program.

The **docs** directory contains user guides, Java API documents, and design documents describing in detail the use of Lexical tools. This document is a general introduction to the programs in the lexical variant generation package for the 2004 version.

The compressed lexical programs are as follows:

lvg2004.tgz

> - The official 2004 distribution of LVG. This includes the source code for the programs, the data and tables in a pure Java embedded database (Instant DB) the programs use, full documentation, installation instructions, and jar files of the programs. See the documents contained within this distribution for a more complete description

of this product.

Normalization (norm)

The lexical program **norm** generates the normalized strings that are used in the normalized string index, MRXNS. Thus norm must be used before MRXNS can be searched.

The normalization process involves stripping possessives, replacing punctuation with spaces, removing stop words, lower-casing each word, breaking a string into its constituent words, and sorting the words in alphabetic order. The uninflected forms are generated using the SPECIALIST lexicon if words appear in the lexicon, otherwise they are generated algorithmically. When a form could be an inflection of more than one base form, the new normalization process returns multiple uninflected forms. If a string to be normalized contains multiple ambiguous forms, and the permutation of these ambiguous forms offer more than 10 output forms, the input form lowercased, with punctuation replaced, word order sorted, but not uninflected, is returned. The upper limit of permutation number (10) is configurable by modifying the configuration file. The program **luiNorm** has the behavior of prior year's normalization, and is distributed for those who need it.

Norm reads its standard input and writes to standard output. It expects input lines to be records separated into fields. The field separator is |. The string to be normalized is identified to norm using the **-t** option. **-t** takes a numerical argument which denotes the field in which the input string is to be found. If no **-t** option appears, norm assumes that the input string is in the first field (**-t:1**). There need not be more than one field, so lines consisting only of input strings are properly understood.

Norm output records include all the fields of the input record with an additional field to the right containing the normalized form of the input string.

For example, if the user had a list of terms to be looked up via the normalized string index in a file called **terms**, he or she could use **norm -i:terms -o:terms.nrm** to get the normalized form of each term. If the input file **terms** contained the following:

    2, 4-Dichlorophenoxyacetic acid
    Syndrome, anterior, compartment
    Abnormal, weight, gain
    Anemia, Refractory, with Excess of Blasts
    left atriums

the file **term.nrm** would contain:

    2, 4-Dichlorophenoxyacetic acid|2 4 acid dichlorophenoxyacetic
    Syndrome, anterior, compartment|anterior compartment syndrome
    Abnormal, weight, gain|abnormal gain weight
    Anemia, Refractory, with Excess of Blasts|anemia blast excess refractory
    left atriums|atrium left
    left atriums|atrium leave

The string in the second field of each line of terms.nrm is now suitable for matching to MRXNS.

Word Index (wordInd)

The lexical program **wordInd** breaks strings into words for use with the word index in MRXW. Users of the word index should use wordInd to break strings into words before searching in the word index. This assures congruence between the words to be looked up and the word index.

Word for this purpose is defined as a token containing only alphanumeric characters with length one or greater. The wordInd program lowercases the output words.

The wordInd program reads its standard input and writes to its standard output. Like norm and lvg, it expects each input line to be a record separated into fields by |. The field containing the input string is identified using the **-t** option. The numerical argument of **-t** denotes the field in which the input string may be found. If no **-t** option is given, the input string is expected to be in the first field (**-t:1**). There need not be more than one field, so lines consisting only of input strings are properly understood.

The wordInd program outputs one line of output for each word found in the input string. Input fields are not repeated in the output unless specified in a **-F** option. Applying wordInd to the input string **Heart Disease, Acute** would result in three output lines:

> heart
> disease
> acute

The numerical argument of **-F** indicates an input field to be repeated in the output. A numerical argument for **-F** option is required for each input field that is to be repeated. Fields are repeated in the order in which the numerical argument of **-F** options appear. The output words always appear as an additional field to the right of any repeated input fields. For example, applying **wordInd -t:2 -F:2:1** to a record of the form **UI23456|tooth, canine|definition.....;** would result in the following output:

> tooth, canine|UI23456|tooth
> tooth, canine|UI23456|canine

The third field of each of those records contains a word extracted from the input term in the first field (**-t:2 ,-F:2**). The **-F:1** option repeats the UI numbers from the first field of input. The fact that **-F:2:1** placed the UI numbers (field 1) after the input string (field 2).

Lexical Variant Generation (lvg)

The lvg program generates lexical variants of input words. It consists of several different flow components that can be combined in various ways to produce lexical variants. The user of lvg chooses combinations of flow components and combines them into a **flow**. (The normalizer program, norm, is essentially the lvg program with a pre-selected flow option: **lvg -f:N**.) The arguments of the **-f** flag are used to specify a flow. Each flow can be thought of as a pipeline with each flow component feeding the next. For example, the flow **-f:i** simply generates inflectional variants and **-f:l:i** generates lowercase inflectional variants. Each of the flow components options is discussed on the documents for lvg.

The lvg program reads from its standard input and writes to its standard output. Input records may be typed in at the keyboard, after typing the command on the command line (**lvg -f:i**) or input lines may be read from a file (**lvg -f:i -i:file**) or piped to lvg from another command (**COMMAND|lvg -f:i**). Output records may be directed to the screen (default), send to a file (**lvg -f:i -i:INFILE -o:OUTFILE**) or piped to another command (**lvg -f:i -i:infile | COMMAND**).

Input

The lvg program is designed to work with one line input records divided into fields. The default field separator is |. The field separator can be changed using the **-s** option. The field in which the input term, whose variants are to be generated, can be specified with the **-t** option. In the absence of a **-t** flag the input term is assumed to be in the first field of the input. So both **dog** and **dog|canine|UI4567** would generate variants of **dog**. With the **-t** flag set to **2**, **dog|canine|UI4567** would generate variants of **canine**. In the case of single field input (**dog**), lvg generates variants from the only field regardless of the setting of **-t**.

The lvg program can read category (part of speech) and inflection information from the input record. The numerical argument to the **-cf** option indicates the field in which category information is located. In the input record, category information needs to be encoded as a number according to the scheme described on the documents for lvg. The numerical argument to the **-if** option indicates the field in which inflection information is located. In the input record, inflection information needs to be encoded as a number according to the scheme described on the documents for Lexical tools.

Output

The lvg program adds five new fields to the input record and outputs a record for each variant generated. For example, if **dog|canine|UI4567** is given to the standard input of **lvg -f:i** the output sent to standard out will be:

> dog|canine|UI4567|dog|128|1|i|1| dog|canine|UI4567|dog|128|512|i|1| dog|canine|UI4567|dogs|128|8|i|1| dog|canine|UI4567|dog|1024|1|i|1| dog|canine|UI4567|dog|1024|262144|i|1| dog|canine|UI4567|dog|1024|1024|i|1| dog|canine|UI4567|dogs|1024|128|i|1| dog|canine|UI4567|dogged|1024|64|i|1| dog|canine|UI4567|dogged|1024|32|i|1| dog|canine|UI4567|dogging|1024|16|i|1|

The first three fields of each record above are identical to the input record, the rest are supplied by lvg. The first additional field is the variant form lvg has generated. The second additional field is the syntactic category of the variant encoded as a number. The third additional field is the inflection of the variant encoded as a number. The fourth additional field indicates the flow that was selected. The fifth field is the number of the flow which generated this variant. Output category (parts of speech) and inflection information are encoded in the same scheme used for input category and inflection information.

Further description of the SPECIALIST lexical tools is available at the SPECIALIST Lexical Tools Website: http://SPECIALIST.nlm.nih.gov/.

[Table of Contents](#)

<div align="center">

Section 5
USING THE UMLS® (UMLSKS) KNOWLEDGE SOURCE SERVER VIA THE INTERNET

</div>

5.0 Background

The UMLS Knowledge Source Server (UMLSKS) is a computer application that provides Internet access to the information stored in the UMLS Knowledge Sources. The purpose of the Knowledge Source Server is to make the UMLS data more accessible to users, and in particular to systems developers. The system architecture is based on the client server model, allowing remote site users (individuals as well as computer programs) to send requests to a centrally managed server at the U. S. National Library of Medicine. Access to the system is provided through a command line interface, through the World Wide Web, an Extensible Markup Language (XML)-based socket programming interface, and through an Application Programmer Interface (API).

Users are encouraged to consult the [UMLSKS Web site](#) for the most current UMLSKS documentation, including the User's Guide, Developer's Guide, and information on downloading the UMLS release files.

5.1 Downloading UMLS Knowledge Source

UMLS Licensees access the UMLSKS and create an account with a login id and password of their choosing. They then use that login id and password for subsequent accesses. Licensees can download the current UMLS Knowledge Sources by going to "Downloads" under "UMLS Knowledge Source" on the server. Downloading is also available under "Resources". Archives of UMLS releases are kept and made available for several previous years on the server. For detailed technical specifications and installation instructions refer to the README.TXT file available on the server.

5.2 System Architecture

The UMLSKS, made available in March 2002, was a redesign of the original "C" programming language system with new features added, including access to the UMLS Knowledge Sources through a public Web interface, incorporation of XML support for programmers both in requesting and returning data, inclusion of a Java-based Object Model of the UMLS Metathesaurus data, and incorporation of a TCP socket-based interface for non-Java programs. Subsequent releases of the

software have augmented the available API functions and refined system operations. The system was designed with following design tenets in mind:

- Extensibility for ease of new feature incorporation
- Scalability in handling ever increasing user loads and increasing numbers of UMLS vocabularies
- Performance considerations permitting faster access to UMLS data
- Flexibility in access modes including a rich API set with access to all of the UMLS data
- Ease of administration by NLM staff and contractors
- Limited system interruptions during system software upgrades

5.3 Querying the Knowledge Source Server

5.3.1 Metathesaurus

The Knowledge Source Server allows the user to request information about particular Metathesaurus concepts, including attributes such as the concept's definition, its semantic types, the concepts that are related to it, etc. It also allows the user to request information about the attributes themselves; for example, by asking for all the concepts that have been assigned to a particular semantic type.

Basic concept information includes the Metathesaurus unique identifier of the concept, the preferred name for the concept, and the names and sources of all terms that comprise that concept. Additional concept information often includes a definition and the source of that definition. Semantic type information is also included. Information about the hierarchical contexts of Metathesaurus concepts is readily available in the system. Related concepts are easily found. If a user were interested in information about a particular term within a concept, then the results could be limited in that way. Co-occurrence data are included for MeSH and AI-RHEUM terminology.

An important perspective on the Metathesaurus is source specific data. It is possible to query the server by limiting the query to a particular vocabulary. The user may wish to see the ancestors or descendants for a term in just a particular vocabulary, or the user may wish to see just the synonyms for a particular term in a particular vocabulary.

Attributes may be queried in the system. Thus, all concepts with a particular semantic type, all terms with a particular syntactic category, and all terms from a particular source vocabulary may be found. Searching for all concepts with a particular semantic type will, for example, give the user a good idea of the coverage of the Metathesaurus in a subject domain.

5.3.2 Semantic Network

The Semantic Network contains information about semantic types and their relationships. The implementation of the network module computes the relationships between semantic types using the inheritance property of the network type hierarchy. Information in the Semantic Network can be queried for semantic types and the relationships between them. Individual queries are specified by providing the known types or relations and leaving out the unknowns. The system then retrieves the corresponding values for the unknowns. For example, if the user wished to know what semantic types are related by a particular relation, then the user would indicate only the relationship name and all the semantic type pairs linked by that relationship would be retrieved. The user might also wish to know if a particular relationship holds between a pair of types.

It is possible to retrieve all the relations between a pair of types. For example, "treats", "prevents", and "complicates" would be listed, among others, as potential relationships between drugs and diseases. It is also possible to retrieve an exhaustive list of all related types in the network. Queries can be made about the definition, unique identifier, tree number, ancestors, parents, children, descendants, and siblings of a semantic type or relation.

5.3.3 SPECIALIST Lexicon

The Knowledge Source Server provides access to lexical records in the SPECIALIST lexicon. The SPECIALIST lexicon is an English language lexicon containing many biomedical terms. The lexicon entry for each word or term records syntactic, morphological, and orthographic information. Lexical entries may be single or multiword terms. Lexical information includes syntactic category, inflectional variation (e.g., singular and plural for nouns, the conjugations of verbs, the positive,

comparative, and superlative for adjectives and adverbs), and allowable complementation patterns (i.e., the objects and other arguments that verbs, nouns, and adjectives can take).

5.4 Gaining Access to the UMLS Knowledge Source Server

Access to the UMLS Knowledge Source Server is available to anyone who has signed the UMLS license agreement and received a license number from NLM. The URL for the Knowledge Source Server Web site is http://umlsks.nlm.nih.gov. First time users should establish a login and a password through the online registration at the web site. Any questions or problems should be addressed via email to umlsks@nlm.nih.gov

5.5 UMLS Knowledge Source Server Documentation

"UMLSKS users should always consult the documentation on the UMLSKS Web site for the most current information."

Additional information on system design, the Application Programmer Interfaces, and future developments is available on the Knowledge Source Server Web site under "About the UMLSKS - Overview".

Also, the following are publicly available on the site, under Documentation:

User's Guide -- describes the basic features of the Web interface, how to navigate the site, and includes information for developers about the two Application Programmer Interfaces.

Developer's API -- documentation generated using the javadoc facility that includes the object model, interfaces and some examples.

A link back to this UMLS Documentation.

Section 6
Metamorphosys - the UMLS install and customization program

6.0 Introduction

MetamorphoSys is the UMLS installation wizard and Metathesaurus customization tool included in each UMLS release. It installs one or more of the UMLS Knowledge Sources; if the Metathesaurus is selected, it allows the user to create customized Metathesaurus subsets.

Users customize their Metathesaurus subsets for two main purposes:

1. to exclude vocabularies from output that are not required or licensed for use in a local application.
   The Metathesaurus consists of a number of files, some of which can be extremely large; excluding sources can significantly reduce the size of the output subset. Given the number and variety of vocabularies reflected in the Metathesaurus, it is unlikely that any user would require all, or even most, of its more than 100 vocabularies. In addition, some sources require separate license agreements for specific uses, which a UMLS user may not wish to obtain. These are clearly indicated in the Appendix to the License Agreement at: http://www.nlm.nih.gov/research/umls/license_appendix.html
2. to customize a subset using a variety of data output options and filters described below.

To identify vocabularies that may not be needed in a customized subset, read the License Agreement and its Appendix, and refer to Appendix B.4 in this documentation. Additional information about some source vocabularies may be found on the

UMLS Homepage under "Metathesaurus Source Vocabularies" at http://umlsinfo.nlm.nih.gov

6.1 MetamorphoSys Requirements

MetamorphoSys has been tested on the following Operating Systems: Sun Solaris (Solaris 8 and 9), Windows (XP, NT and 2000), Linux and Macintosh (OS x 10.3 with current JRE from Software Update). It is implemented in Java and run-time environments for all platforms included in the release.

You may execute MetamorphoSys from the UMLS DVD-ROM (which contains the application and the compressed UMLS Knowledge Sources data files) or use a high-speed Internet connection to download files from the UMLS Knowledge Sources Server (UMLSKS) at: http://umlsks.nlm.nih.gov. Because downloads on a T1 line with 1 megabit per second throughput will require over 5 hours, we expect that MetamorphoSys will usually be run from the DVD.

To use the DVD, you must have a DVD reader and at least 26 GB free space. Multiple runs that create multiple subsets of the Metathesaurus will need even more space. For reasonable performance, we suggest these minimum requirements:

- A fast CPU - 1GHz or higher
- 6x (or better) DVD drive
- 1 GB of RAM, preferably more

DVD options allow the user to (1) copy UMLS Knowledge Sources data files to local storage and (2) copy MetamorphoSys .nlm data format files to local storage. This may be useful for multiple runs or subsetting an existing subset, and it may improve performance time.

A third option, Validate Distribution, allows users to verify the integrity of .nlm files downloaded from the UMLSKS or copied from the UMLS DVD. It compares downloaded file sizes to those in the release .CHK file, and is a useful first step for trouble-shooting when problems occur with a UMLS installation.

If the UMLS release is downloaded from the UMLSKS, there are five associated files:

- mmsys.zip (zipped MetamorphoSys application)
- 2005AC-1-meta.nlm (compressed Metathesaurus data)
- 2005AC-2-meta.nlm (compressed Metathesaurus data)
- 2005AC-3-meta.nlm (compressed Metathesaurus data)
- 2005AC-otherks.nlm (compressed Semantic Network and SPECIALIST Lexicon)

The mmsys.zip file is first unzipped to local storage and the MetamorphoSys application started.

6.2 Starting MetamorphoSys

Open a terminal window and change to the root directory of the DVD-ROM. Type the appropriate command for your platform:

- MACINTOSH   ./macintosh_mmsys.sh
- LINUX   ./linuxmmsys.sh
- Solaris   ./solaris_mmsys.sh

Press the return key

A new window will appear. Be patient since a good deal of software must load before the Welcome screen appears.

- Windows   windows_mmsys.bat

On Windows machines with Autorun enabled, the DVD will start automatically. If it does not, go to the root directory of the

DVD-ROM and click on the file named windows_mmsys.bat.

6.3 Using MetamorphoSys

MetamorphoSys screens and tabs will lead you through the process of installing all the UMLS Knowledge Sources and customizing the Metathesaurus.

6.3.1 Welcome to MetamorphoSys

Select either:

Install UMLS	to install one or more UMLS Knowledge Sources
OR	
Customize My Subset (See Section 6.5)	to further customize an existing Metathesaurus subset
OR	
Browse my Subset	to open the RRF Subset browser
OR	
Build MRCXT	to open the MRCXT Builder

Other options located on the File menu include "Validate Distribution" and "Copy to Hard Drive".

Validate Distribution confirms that all UMLS files, whether they have been downloaded or distributed via DVD, have transferred correctly and are complete. The process takes approximately 30 minutes, and produces a log file (validation.log) and an alert box displays a statement regarding the validity of the files. Use "Validate Distribution" as your first step in troubleshooting when experiencing any malfunctions.

Copy to Hard Drive copies the Knowledge Sources data files, identified by the ".nlm" extension, and MetamorphoSys to local storage. Local storage may improve startup times for MetamorphoSys as contrasted to running off the DVD.

6.3.1.1 Install UMLS

MetamorphoSys creates a top-level destination directory in local storage for the UMLS Knowledge Sources. The directory is named with the release version, e.g., 2005AC. The following directory structure is created beneath the destination directory, shown below for the 2005AC release:

<installation directory>

        2005AC
                NET
                LEX
                META
                MMSYS

You may install any one, two, or all three Knowledge Sources. When selected,

- the Semantic Network is installed to the NET directory

- the SPECIALIST Lexicon is installed to the LEX directory

- MetamorphoSys is installed to the MMSYS directory

The META directory is populated with the Metathesaurus subset files created during installation, depending on your configuration some of these files may contain zero bytes.

Use the Browse button to locate source and destination directory locations.

Click OK to proceed with installation. A Progress monitor tracks each step of the Installation process. If the Metathesaurus is selected installation will begin after all Metathesaurus options are selected.

Cancel the installation at any time. Click Cancel at the bottom of the Install UMLS progress screen, or at the bottom of the MetamorphoSys Progress window.

6.3.1.2 MetamorhpoSys Configuration

Select 'New Configuration' to create a new subset configuration or if you have not previously used MetamorphoSys. If you previously installed the UMLS Knowledge Sources and created a Metathesaurus configuration file, select "Open Configuration" to utilize that file.

6.3.1.2.1 License Agreement Notice

The Metathesaurus contains source vocabularies produced by many different copyright holders. The majority of the content of the Metathesaurus is available for use under the basic (and quite open) terms described in the Metathesaurus license at: http://nlm.nih.gov/research/umls/license.html

However, some vocabulary producers place ADDITIONAL RESTRICTIONS ON THE USE OF THEIR CONTENT AS DISTRIBUTED WITHIN THE METATHESAURUS.

Three levels of additional restrictions are described in Section 12 of the license. Individual vocabularies and their restriction levels are listed in the Appendix to the UMLS License Agreement. If a user already has a separate license for use of one of the source vocabularies, the existing license also applies to that source as distributed within the Metathesaurus. In some cases, UMLS users may have to request permission or negotiate a separate license with a vocabulary producer in order to use that vocabulary in a production system. There may be a charge associated with these separate permissions or license agreements.

Click "Accept" or "Do Not Accept" after reviewing the license agreement.

6.4 Select Default Subset

Two default subsets have been defined for creating useful and manageable output subsets. Others may be added in the future based on user feedback. During initial installation of the Metathesaurus, you must select one of two default subsets as a starting point:

1. Level 0 - contains vocabulary sources for which no separate, additional license agreements are necessary beyond the UMLS license.

2. Level 0 + SNOMED CT - contains all Level 0 sources (no additional licenses needed for sources) and SNOMED CT.

Note: Non-U.S. users must have separate license agreements to use SNOMED CT (see Section 12 in the UMLS license agreement at: http://www.nlm.nih.gov/research/umls/license.html

You will have the opportunity to modify your default subset to include or exclude additional sources using the Source List tab (see Section 6.6.3 below).

Please note: The default RxNorm Subset, which contained RxNorm concepts in Level 0 sources, has been removed. Users can produce the same subset by applying the RxNorm Filter. Access this filter using the File Menu, Enable/Disable Filters option. An additional MetamorphoSys tab for the RxNorm filter will appear when the filter has been selected. For more information regarding the RxNorm Filter see section 6.6.1.5.

6.5 Option Tabs

Five basic Options Tabs--Input Options, Output Options, Source List, Precedence, and Suppressibility--provide a variety of customization options.

You may select and complete Option tabs in any order. Note that the selections that you make in one option may affect the data displayed, and the choices available, on other Options tabs.

You may return to the default settings for any option. Select Reset on the menu bar, and select the appropriate Reset command.

When you have completed configuring your Metathesaurus subset, go to the menu bar; select Done, and Begin Subset.

You will be prompted to save your configuration. Name your configuration file, which will be stored in the destination META directory. This file documents your configuration choices, and can be used as the starting point for a later customization using the Customize My Subset option on the Welcome screen.

6.5.1. Input Options

Allows users to indicate the location of required directories, the configuration file, and the input and output directories.

For the initial installation, NLM Data File Format must be selected.

If you are customizing an existing subset, use Browse to select its current format of either Original Release Format or Rich Release Format.

You may select and complete Option tabs in any order. Note that the selections that you make in one option may affect the data displayed, and the choices available, on other Options tabs.

You may return to the default settings for any option. Select Reset on the menu bar, and select the appropriate Reset command.

When you have completed configuring your Metathesaurus subset, go to the menu bar, select Done, and Begin Subset.

6.5.2. Output Options

6.5.2.1 Select Output Format

Select either Original Release Format or Rich Release Format. Rich Release Format is the default selection for the initial installation and for customizing an existing subset in the Rich Release Format. Original Release Format is the default for customizing an existing subset in the Original Release Format.

Note: You cannot generate a correct Rich Release Format subset from Original Release Format.

6.5.2.2 Subset Folder

Indicate where the new subset files should be placed.

6.5.2.3 Write Database Load Scripts

Outputs a load script in either Oracle or MySQL format, which you may further optimize or customize. For more information on UMLS load scripts go to: http://www.nlm.nih.gov/research/umls/load_scripts.htmll

6.5.2.4 Source Abbreviation Format

Source vocabulary information in the Metathesaurus content can be identified by a versionless, or Root Source Abbreviation (RSAB), or by the longer and more descriptive Versioned Source Abbreviation (VSAB). The default is the RSAB, but you may choose to include the VSABs. For example,

MSH	Root Source Abbreviation (RSAB)
MSH_2003_12_12	Versioned Source Abbreviation (VSAB)

In either case, your subset will include the MRSAB file which links the RSABS to the corresponding VSABs for all source vocabularies in your subset.

6.5.2.5 Maximum Field Length

Restrict fields in your output to the maximum field length allowed in your application or database software.

6.5.2.6 Eliminate Extended Unicode Characters:

This option allows users to select output encoding in either 7-bit ASCII or UTF-8. 7-bit ASCII is the default output from MetamorphoSys. Select this box to output data in Unicode UTF-8 format. See 6.5.2.5.1, UMLS Character Sets, in the next section.

You may select and complete Option tabs in any order. Note that the selections that you make in one option may affect the data displayed, and the choices available, on other Options tabs.

You may return to the default settings for any option. Select Reset on the menu bar, and select the appropriate Reset command.

When you have completed configuring your Metathesaurus subset, go to the menu bar; select Done, and Begin Subset.

6.5.2.6.1 UMLS Character Sets

Previous releases of the Metathesaurus used a simple 'least common denominator' character set known as '7-bit ASCII' or 'Basic Latin.' This standard contains no diacritics or special symbols, and is the default output encoding from MetamorphoSys.

Beginning with the 2004AA release, extended characters (including diacritical marks, ideographs, and scientific and other symbols) are also supported in Unicode, specifically in the the UTF-8 format of the Unicode 4.0 standard [1]. Unicode is the emerging international data encoding standard, and currently represents 96,382 different characters from the world's scripts and most languages.

Users may elect to output extended characters in MetamorphoSys. Further, users may choose to convert UTF-8 to other character sets using tools and online data tables available at http://www.unicode.org. The UMLS does not include character set conversion tools.

When extended characters appear in a source string, they are converted to UTF-8 in the Metathesaurus as necessary. For English sources, i.e., LAT =ENG, an equivalent 7-bit ASCII string is also created for the UMLS using the lvg program (see http://umlslex.nlm.nih.gov) to ensure that no information is lost when using the 7-bit ASCII character set.

1) 7-bit ASCII or Basic Latin (used in Original Release Format)

This is the 'least common denominator' character set of 96 characters and symbols from the original ASCII standard. The UMLS 7-bit ASCII characters include those from 32 decimal to 127 decimal and these "C0 Controls" shown below:

0x09 CHARACTER TABULATION, horizontal tabulation. HT

0x0A LINE FEED, new line (NL), end of line (EOL)

0x0D CARRIAGE RETURN

This is the character set used in the original 'MR' format (Original Release Format) files, compatible with most computer systems.

2) Unicode and UTF-8 (optional extended characters; used in Rich Release Format, beginning with the UMLS 2004AA Release)

Only users who are familiar with Unicode and have fully compliant systems should enable the UTF-8 output of extended characters in MetamorphoSys.

Data from some sources in the Metathesaurus is now encoded in UTF-8 (UTF=Unicode Transformational Format), an 8-bit encoding suitable for processing in byte-oriented computer systems. It is a variable length encoding, so that a character can span one or more bytes. The initial byte order mark (BOM) character is not present in the UTF-8 encoded Metathesaurus files.

Note that UTF-8 is identical to the ASCII encoding for characters in the 7-bit ASCII range, so that 7-bit ASCII files are automatically a correct subset of UTF-8; this subset is the MetamorphoSys default.

The NLM receives data in a variety of character sets from source providers. Typically, files are encoded in variations of ASCII, including the ISO 8859 or the Windows codepage 1250 families of character sets. When not supplied in UTF-8, NLM will convert new or updated Sources to UTF-8, as resources permit.

The Metathesaurus has historically contained names for concepts from languages other than English (the different translations of the MeSH vocabulary, for example), but the characters used to represent these names were coerced to ASCII using a transliteration scheme. When users need correct extended characters and when non-European sources are added to the UMLS, the information loss becomes unacceptable and transliteration may be completely impossible.

From the 2004AA release forward:

1. Names of concepts in all vocabularies will be represented as supplied but converted to Unicode when necessary.
2. A 7-bit transliteration of strings in Western European languages that contain extended characters will be added, using the new lvg flow (see below). These strings will be identified in the MRCONSO.RRF file with the SAB and TTY columns set to **INSERT SABs* to select or exclude them.
3. There will be a relationship linking the extended character strings and their 7-bit translations.
4. Files will be in byte sort order (with data in UTF-8, standard UNIX sort works as expected). Note that the UMLS data are intended to be manipulated with software tools such as database systems, so the sort order of the files should not matter.

6.5.2.6.2 LVG flow

LVG stands for "lexical variant generation" and is a set of tools and data that are distributed with the UMLS as part of the SPECIALIST system (see http://umlslex.nlm.nih.gov/). The current version of LVG includes flows to convert UTF-8 strings into a canonical 7-bit representation that includes the removal of diacritics, expansion of ligatures and the substitution of official Unicode character names with appropriate escape character sequences for the remaining Unicode characters [2].

6.5.2.6.3 MetamorphoSys Support

MetamorphoSys is Unicode compliant. By default, it will eliminate rows that contain extended characters (those not in the 7-bit ASCII range). Note that some English language sources may contain Unicode characters in names and attributes.

6.5.2.6.4 OS and Database Support

Most modern Operating Systems are Unicode (and UTF-8)-aware. For example, Solaris 2.9, Windows XP, and most Linux systems can store, process, and display information that is encoded in UTF-8, though the task of migration may not necessarily be painless.

Database vendors are also starting to migrate to UTF-8, but understandably often lag the OS vendors. Oracle and MySQL (version 4.1 and up [3]) in our experience seem to work correctly.

Third party software may not always work correctly with Unicode data. Check with your vendor or software provider.

6.5.2.6.5 References

1. The Unicode Standard 4.0, Unicode Consortium, Addison-Wesley, http://unicode.org
2. Lexical Variant Generation http://umlslex.nlm.nih.gov
3. MySQL Documentation, Chapter 9, National Character Sets and Unicode, http://www.mysql.com/doc/en/index.html

6.5.2.7 Exclude MRCXT (MRCXT.RRF)

Allows users to exclude the very large MRCXT or MRCXT.RRF file from their output, reducing MetamorphoSys processing time, and significantly reducing the size of the resulting subset.

If Original Release Format is the selected output format, MRCXT can be excluded. If Rich Release Format is selected, MRCXT.RRF can be excluded.

See Section 2.7.1.3.11 for information about how MRHIER.RRF can be used to compute hierarchies.

6.5.2.8 Remove MTH only concepts

Select this option to retain MTH atoms ONLY when they overlap with atoms from other sources in your subset.

6.5.2.9 Calculate MD5 values for output files

When this box is checked, the MD5 algorithm is used to generate a "mmsys.md5" file in the Metathesaurus subset directory. The information in this file can be used to verify data integrity of the Metathesaurus files (RRF or ORF), and can be useful when troubleshooting problems. The MD5s values appear in the META/mmsys.md5 file. Please note that these MD5s are intended for comparison of different runs and are calculated in a platform independent way, i.e., they ignore differing line terminations. For this reason, native MD5 calculation programs may differ from those in the mmsys.md5 file.

6.5.2.10 Add UTF-8 BOM characters to output files

When this box is checked, all output data files are prepended with a byte order mark. This beginning-of-file marker (3 bytes) indicates that the file is encoded as UTF-8.

6.5.3 Source List

The Source List tab displays all source vocabularies in the current version of the Metathesaurus. Sources are sorted alphabetically by Source Abbreviation in the default display. At the top of the Source List tab there are two radio buttons:

source list options

The highlighted sources reflect the default subset selected earlier in the installation process. You may select or deselect additional sources to include or exclude from your subset. Leave the button set to "Select Sources to EXCLUDE from subset in order to highlight sources that will be removed from your customized Metathesaurus subset.

Or, you may choose "Select sources to INCLUDE in subset ." When selected, only the highlighted sources will be included in your local subset.

Note: The highlighted sources do NOT change when you switch between these two options. If a source is highlighted for EXCLUSION from a subset, and you choose "Select sources to INCLUDE in subset ," that source will now be highlighted for INCLUSION in your subset.

To select or deselect additional rows, hold down the <CTRL> key while making your selection.

You may sort the Source List by Full Source Name, Source Abbreviation, Source Family, Language or Level (UMLS License Restriction Level). Click on the column header to resort the list by that data.

The complete Metathesaurus contains over 100 source vocabularies and in its entirety is an extremely large and unwieldy set of data files. Carefully consider what sources will contribute useful data to your application, and then exclude other sources, to reduce the size of output subsets and to improve application performance.

Consider also that the data from some sources may be incompatible with your intended application. They may contain terms that are recognizable only within the context of a specific source; or they may contain abbreviations that are confusing, or not particularly useful to your application.

Additional information on a few specific sources is available under "Metathesaurus Source Vocabularies" at: http://www.nlm. nih.gov/research/umls. Users may also contact the source providers included in the Appendix to the License Agreement for additional documentation or information.

You may select individual sources to remove based on the Full Source Name or Source Abbreviation. You may take advantage of groups of related vocabularies, called Source Families, to assist in the removal of related sources when one source is selected.

Note, for example, that CPT (the AMA's Physicians' Current Procedural Terminology, CPT4) is also a part of HCPT (the Health Care Financing Administration Common Procedure Coding System, HCPCS). Both vocabularies must be removed to exclude all sources of CPT information.

You may also exclude sources by language, or by license restriction level. To reset source selections and return to the default list, select "Reset Sources to Exclude Defaults" under Reset on the menu bar.

You may select and complete Option tabs in any order. Note that the selections that you make in one option may affect the data displayed, and the choices available, on other Options tabs.

You may return to the default settings for any option. Select Reset on the menu bar, and select the appropriate Reset command.

When you have completed configuring your Metathesaurus subset, go to the menu bar; select Done, and Begin Subset.

6.5.4 Precedence

The Precedence tab displays the default order of precedence of Metathesaurus source and term type combinations as determined by NLM. One string from one English term is designated and labeled as the default preferred name of each concept in the Metathesaurus. Selection of the default preferred name for any Metathesaurus concept is based on an order of precedence of all the types of English strings in all the Metathesaurus source vocabularies. Different types of strings, e.g., preferred terms, cross references, abbreviations, from each vocabulary will have different positions in this order.

The default order of precedence determined by NLM will not be suitable for all applications of the Metathesaurus. MetamorphoSys can be used to change the selection of preferred names to feature terminology from the source vocabularies most appropriate to particular user populations.

A user may reorder the ranking of source and term type combinations by cutting and pasting, or dragging and dropping, the rows in the Precedence List. Term types from sources that have been excluded on the Source List tab will not be displayed.

Shift rows by cutting and pasting the rows. Multiple rows can be cut by holding the <CTRL> key down while making selections. To paste the rows, select the location where the rows will be pasted and press <CTRL-V>.

The ranking of sources and term types will affect the output subset. In particular, the name of a concept will be determined by the highest ranking term type in that concept.

You may select and complete Option tabs in any order. Note that the selections that you make in one option may affect the data displayed, and the choices available, on other Options tabs.

You may return to the default settings for any option. Select Reset on the menu bar, and select the appropriate Reset command

When you have completed configuring your Metathesaurus subset, go to the menu bar, select Done, and then Begin Subset.

6.5.5 Suppressibility

The Suppressibility tab displays source/term type combinations to be marked as suppressible in the output subset. Term types from sources that have been excluded on the Source List will not display. For a new subset, the initial display highlights default source/term types made suppressible by NLM. Users may select or deselect source/term types to be marked as suppressible in their output subsets. When customizing an existing subset, the initial display highlights the user's suppressibility settings for that subset.

You may select and complete Option tabs in any order. Note that the selections that you make in one option may affect the data displayed, and the choices available, on other Options tabs.

You may return to the default settings for any option. Select Reset on the menu bar, and select the appropriate Reset command.

When you have completed configuring your Metathesaurus subset, go to the menu bar; select done, and Begin Subset.

In addition to the 5 basic option tabs, there are additional filters that can be used to customize the Metathesaurus. See 6.7.1.2, Enable/Disable Filter.

6. 6 File Menu

6.6.1. Enable/Disable Filter

This command allows users to enable any one or all of six (6) additional filters: Attributes To Exclude, Languages To Exclude, Relationships To Exclude, Semantic Types To Exclude, RxNorm, and content View. When these filters are enabled, additional tabs appear on the UMLS Metathesaurus Configuration screen. When one of these filters is disabled, its tab disappears.

6.6.1.1 The Attributes To Exclude Filter allows users To Exclude specified attributes from their output subset. Use <CTRL> mouse click to (de)select source attributes from the table. This filter removes only attribute data and not entire concepts from the output subset. Term types from sources that have been excluded on the Source List tab will not be included on the Attributes To Exclude tab. When configuring this filter, (de)selecting an attribute type for removal will prompt the user with a list of all other attributes from the same source so they can considered and (de)selected if appropriate. The default list is sorted alphabetically by Source Name. Users may sort the list by Source Abbreviation or Attribute Type by clicking on the column header.

6.6.1.2 The Languages To Exclude Filter allows users To Exclude specific languages from the output subset. All terms with the specified languages will be removed as well as all attributes and relationships connected to those terms. If every term in a concept is in a language selected for exclusion, then the entire concept will be removed.

6.6.1.3 The Relationships To Exclude Filter allows users to specify relationship types To Exclude from the output subset. Use <CTRL> mouse click to (de)select source relationships from the table. This filter removes only relationship data and not entire concepts from the output subset. Relationship from sources that have been excluded on the Source List tab will not appear on this option tab. When a relationship type has been selected for removal, MetamorphoSys will display a list of all other relationships with the same source, so that they can also be considered and (de)selected if appropriate. The default list is sorted alphabetically by Source Name. Users may sort the list by Source Abbreviation or Relationship Type by clicking on the column header.

6.6.1.4 The Semantic Types To Exclude Filter allows users to specify a list of semantic types to be used for concept removal. The default behavior is to remove any concepts from the output subset containing at least one semantic type from the specified list. When configuring this filter, selecting a semantic type for removal will cause MetamorphoSys to prompt the user with a list of all children of that semantic type so they can also be (de)selected. The default list is sorted by Semantic Hierarchy. The user may sort the list by Semantic Type Unique Identifier (TUI) or Semantic Type by clicking on the column header. When this filter is enabled, Advanced Semantic Types To Exclude Options becomes available under the Options menu. See Section 6.8.4 below.

6.6.1.5 RxNorm Filter

This filter extracts from the Metathesaurus only those concepts that meet one of the following criteria:

a) RxNorm atoms, e.g., SCD, or
b) HL7 dose form atoms (TTY=DF), or
c) Semantic Type = Drug Delivery Device.

RxNorm is a standardized nomenclature for clinical drugs. The RxNorm name of a clinical drug combines its ingredients, strengths, and form in which the drug is administered or is specified to be administered in a prescription or order. RxNorm is produced by the National Library of Medicine, and has been designated a HIPAA standard vocabulary. For more information on RxNorm, see: http://www.nlm.nih.gov/research/umls/rxnorm_mail.htm

The RxNorm filter includes RxNorm concepts from sources at license restriction Level 0 only. RxNorm standard names for clinical drugs are connected to the varying names of drugs present in many different controlled vocabularies within the UMLS Metathesaurus, including commercial drug information sources. You may further customize your RxNorm subset by using the Source List tab to add or remove other sources that may have names for RxNorm concepts.

6.6.1.6 Content View Filter

A content view is any definable subset of the Metathesaurus that is useful for some specific purpose. The actual definition of a content view can take a variety of different forms: 1. An actual list of Metathesaurus UIs, maintained over time. 2. A list of sources that participate in the view. 3. A complex query that identifies particular sets of data.

The first Content View Flag (CVF) to be made available identifies sets of terms that are useful for Natural Language Processing. This CVF represents the strict model of data used by Metamap. These terms carry the value "256" in the CVF field.

6.6.2 Import Filter

This command allows the user to import filters developed according to the Filter API into the application. Filters cannot be exported or removed from the application, but they can be disabled. A window will pop up with all filters available for import. These filters are found in the METAMSYS/ext directory. See Section 6.11 for more information.

Two simple import filters are provided as examples of custom filtering:

      NosNec (for Testing):   To exclude "NOS" or "NEC" strings from the output subset.
      OddEven (for Testing):   To exclude odd or even numbered CUIs from the output subset.

When an import filter is selected, its option tab appears on the Metathesaurus configuration screen.

6.6.3 New Configuration

Use this command to create a new subset configuration. The License Agreement Notice is displayed (see Section 6.3.1.1.1.) and the configuration process continues as described in Section 6.3 and following.

6.6.4 Open Configuration

Use this command to open a previously saved configuration, which can be run (go to Done, and Begin Subset) or modified. MetamorphoSys displays the config directory in the MMSYS folder as a starting point from which to locate and select a previously saved configuration.

6.6.5 Save Configuration

Use this command to save the current configuration. MetamorphoSys prompts the user to assign a file name and displays the top level UMLS directory as a starting point for storing the saved configuration file. This allows a user to save a configuration and run it to produce the Metathesaurus subset at a later time. The saved configuration can also be further modified to create new subset configurations.

6.6.6 Exit

Use this command to exit MetamorphoSys. A prompt provides an opportunity for the user to save the configuration before exiting.

6.7 Edit Menu

Two commands, Increase Font and Decrease Font, allow the user to change the text size displayed in MetamorphoSys screens. An additional command, Undo Enable Filter, is available if any filters have been enabled from the File menu.

6.8 Options (for Advanced Users)

Advanced options include MetamorphoSys Options, Advanced Source List Options, and Advanced Suppressibility Options.

6.8.1 MetamorphoSys Options

Opens a configuration window which contains the following user capability.

Auto Select Related Items - If this check-box is selected, there is no prompt when the selected row shares a Source Family or has a Dependent Source. The system selects the Dependent Source rows or the rows with the same Source Family automatically. The default for this flag is false.

6.8.2 Advanced Source List Options

Opens a configuration window which contains the following user capabilities:

6.8.2.1 Enforce Family Selection - If the "Enforce Family Selection" check-box is selected, the user will be prompted to select other sources in the same "Source Family."

6.8.2.2 Enforce Dependent Source Selection - If "Enforce Dependent Source Selection" is selected, and the user selects a source in the "Dependent Source Associations" table, the user may select any dependent sources listed. As with "Enforce Family Selection" this functionality exists for deselection of sources as well. The default for this flag is true.

Source/Dependent Source relationships can be added to the "Dependent Source Associations" table by clicking on the "Add" button. The user may clear the whole table by clicking on the "Clear" button. A specific line or lines can be removed from the table by selecting those lines and pressing the "Delete" button. The user may also sort the table either by clicking on the "Source" or "Dependent Source" table header. A reverse sort of the table can be done by pressing while clicking on a table header. The user may exit the "Advanced Options" dialog by clicking on the "Done" button at the bottom of the window.

6.8.3 Advanced Suppressibility Options

If the "Remove Suppressible Data" check-box is selected, all data in which

    ts='s' or 'p' (in MRCON; Original Release Format)
                    OR
    SUPPRESS flag is set to Y (in MRCONSO.RRF)

will be removed from the result set.

For example, the following rows would be among those removed:

MRCON (ORF)
C0000731|ENG|s|L0658950|PF|S0835542|Change in abd size/distension|0|

MRCONSO (RRF)
C0000731|ENG|S|L0658950|PF|S0835542|Y|A0894040||||ICPC|PS|D25|Change in abd size/distension|0|Y||

The corresponding rows will be removed from other files containing the same CUI, SUI. If this operation causes all rows for a CUI to be removed in MRCON, that entire CUI will be excluded from the result set for the other files. The default setting for this flag is N (No).

See also Suppressibility, Section 6.5.5 above.

6.8.4 Advanced Semantic Types To Exclude:

These options are available when the Semantic Types To Exclude filter has been enabled from the File menu, and allow the user to set the predicate for concept removal. There are two choices:

1. Remove CUIs containing at least one selected semantic type - If this option is selected, a concept will be removed if any of its semantic types appear on the exclude list.
2. Remove CUIs containing only selected semantic types - If this option is selected, a concept will be removed only if all of its semantic types are on the exclude list.

6.9 Reset Menu

The Reset menu allows the user to return to Metathesaurus default selections for all of the filter tabs (Input Options, Output Options, Source List, Precedence and Suppressibility). The choice of version, Original Release Format or Rich Release Format, will not be reset on the Output Options tab and the Input options tab. The default selections are those listed in the mmsys.prop.default file in the config folder. The mmsys.prop.sav file contains the properties used in the last run of MetamorphoSys.

6.10 Done; Begin Subset

When all options have been explored and you have completed configuring your Metathesaurus subset, select Done from the menu bar, and then Begin Subset. If you would prefer to save your configuration in order to subset at a later time, select "Save Configuration" from the File menu.

The Install UMLS Metathesaurus progress monitor charts the process through the following steps: Initializing the CUI list; Subsetting Content, Subsetting Indexes, and Final Processes. To stop processing and exit MetamorphoSys at any time, press Cancel at the bottom of the progress monitor. The interrupted process cannot be resumed. The configuration must be recalled (if saved), or recreated (if not saved), and subsetting must be started again.

MetamorphoSys produces an "install.log" file in the user's release directory, containing the log of the installation process up to the start of Metathesaurus subsetting. It records which operations were selected, and reports the results of file validations against both CHK and MD5 files. If the downloaded files pass validation, processing continues and subsetting begins. If files fail validation, the install.log is displayed.

When subsetting is complete, progress and errors messages, and the configuration settings, are displayed on the screen and also written to a log file called "mmsys.log" in the directory containing the subsetted files. The subsetted Metathesaurus files are located in the chosen destination directory (see Section 6.3.1.1).

6.11 API Documentation

To help users develop custom filters, the MetamorphoSys API documentation (generated with javadoc) can be found starting with the file METAMSYS/doc/index.html in your installation directory. Sample filters using this API can be found in the METAMSYS/ext directory. Additional filters may become available at http://umlsinfo.nlm.nih.gov. Check the information there, and especially at: http://umlsinfo.nlm.nih.gov/mmsys.

6.12 Version Validation

In order to obtain create correct subsets, the user MUST use the version of MetamorphoSys that matches the version of the Metathesaurus release files being subsetted. Do not use older versions of MetamorphoSys with newer or older release files; use the version of MetamorphoSys included with the release files.

6.13 Getting Help

Check the information available at: http://umlsinfo.nlm.nih.gov. We are developing additional Web resources based on user input.

NLM maintains a listserv (electronic mailing list service) called umls-users where users can share their experiences with, or seek advice from the UMLS community on using UMLS resources. NLM also uses this forum to seek advice from UMLS

users, and to distribute official announcements about UMLS products and services.

To subscribe, send an email to: listserv@list.nih.gov containing the following message: SUBSCRIBE|SIGNUP UMLSUSERS-L <your full name>

To post a message to the list AFTER subscribing, send email to: UMLSUSERS-L@list.nih.gov

To access subscription information and list archives, go to: UMLSUSERS-L Listserv Webpage

An alternative list, UMLS-ANNOUNCES-L, exists for users who wish to receive only official announcements about UMLS products and services, including new releases, new features, and problem/fix messages.

To subscribe, send an email to: listserv@list.nih.gov containing the following message: SUBSCRIBE|SIGNUP UMLS-ANNOUNCES-L <your full name>

To access subscription information and list archives, go to: UMLS-ANNOUCES-L Listserv Webpage

6.14 Acknowledgments

Solaris and Windows Java Runtime Environment: http://javasoft.com

Linux Java Runtime Environment: http://www.blackdown.org

Table of Contents

Section 7
UMLS DVD

7.0 Introduction

The UMLS is available on a single disk DVD, which contains the three UMLS Knowledge Sources in four (4) compressed format files (.nlm data format)

- 2005ac-1-meta.nlm
- 2005ac-2-meta.nlm
- 2005ac-3-meta.nlm (Metathesaurus = 3 files; total 2 GB)
- 2005ac-otherks.nlm (Semantic Network and SPECIALIST Lexicon; 501 MB)
- MetamorphoSys, the UMLS installation and Metathesaurus subsetting program for PC and UNIX machines (180 MB cross-platform Java executable with JRE for each supported platform).

UMLS licensees may request the DVD from: umls_support@nlm.nih.gov and must include their license number in the request.

7.1 Hardware and Software requirements

7.1.1 Supported operating systems:
      Windows: XP, 2000, NT
      Linux*
      Solaris: Solaris 8 and 9
      Macintosh: OS X 10.3 with current JRE from Software Update

*Linux note: The specific version fully validated is Red Hat Enterprise Linux WS, Version 3. Other versions should also work; NLM will post reports of other Linux versions as users report success.

7.1.2 Hardware Requirements

- A MINIMUM 26 GB of free hard disc space.
    Yes, that's over 26,000 megabytes!
- A MINIMUM of 1 GB of RAM, preferably more. Smaller memory size will cause virtual memory paging with exponentially increased processing time.
- A CPU speed of at least 1 GHz for reasonable installation times.
- DVD drive (the faster the better).

7.2 Installing from DVD

Insert the DVD into the DVD drive. For best results the drive should be 6X or higher.

Start the MetamorphoSys install program:

Windows

- The DVD should autorun.
- If not (or if installing from the hard disk), go to the DVD root directory and click on "windows_mmsys.bat."

Linux, Solaris, Macintosh

- Open a terminal window and change to the root directory of the DVD, then type the appropriate command for your platform:
        ./linuxmmsys.sh
        ./solaris_mmsys.sh
        ./macintosh_mmsys.sh
- hit the return key

Be patient. A good deal of software must load before the welcome screen appears.

- Select: Install UMLS on the welcome to MetamorphoSys screen
- Select destination (local) directory for files on the Install UMLS screen

The Install UMLS Metathesaurus progress monitor charts the process through:

- Initializing the CUI list
- Subsetting Content
- Subsetting Indexes
- Final Processes

If selected, Semantic Network and SPECIALIST Lexicon files are copied first. Accept the License Agreement notice to proceed with customizing the Metathesaurus.

Select Cancel to exit MetamorphoSys at any time. The interrupted process cannot be resumed. The configuration must be recalled (if saved), or recreated (if not saved), and subsetting must be started again.

MetamorphoSys produces an initial install.log file of the installation up to the start of Metathesaurus subsetting. If files pass validation, processing continues and subsetting begins. If validation fails, a warning is displayed and recorded in install.log.

When subsetting is complete, configuration settings are displayed on the screen and also written to "mmsys.log" in the directory containing the subsetted files. Your customized Metathesaurus files are located in the destination directory.

For more information on running MetamorphoSys see Section 6 of this documentation.

APPENDIX A

License Agreement for Use of the UMLS® Metathesaurus®

The License Agreement for 2005 is available with the 2005AC Documentation: (license.html) and on the NLM website: (http://www.nlm.nih.gov/research/umls/license.html). NLM does not charge for the UMLS Knowledge Sources. Users of the UMLS Metathesaurus may have to enter into separate license arrangements (See Appendix A.1), which may involve charges, with the copyright holders of some of the individual vocabularies that have been incorporated in the UMLS Metathesaurus.

Send questions, comments about the UMLS project to: custserv@nlm.nih.gov or call 1-888-FINDNLM.

APPENDIX A.1

Appendix to the License Agreement for Use of the UMLS® Metathesaurus

UMLS METATHESAURUS® SOURCE VOCABULARIES - 2005AC Edition

Sources are listed in order according to the abbreviations used in the UMLS Metathesaurus files. If additional restrictions and notices apply, the category of restrictions and the special notices appear under the name of the source. See the license agreement for an explanation of the categories of restrictions. Many sources publish printed editions and/or other explanatory information that may be essential to understanding the purpose and application of particular sources in data creation and retrieval. Contact information is provided for each source. Please address questions about permissions or license agreements for additional uses not covered by this Agreement, or other inquiries about individual sources, to the appropriate contacts.

NLM is working toward inclusion in the UMLS Metathesaurus of the complete, current edition of most of these vocabulary sources.

AIR93 - AI/RHEUM. Bethesda, (MD) : National Library of Medicine, Lister Hill Center, 1993.

Contact: May Cheh, Lister Hill National Center for Biomedical Communications, National Library of Medicine, Building 38A, Room 9E902, 8600 Rockville Pike, Bethesda, MD 20894; e-mail: cheh@nlm.nih.gov.

ALT2005 (updated) - Alternative Billing Concepts (Altlink). Version 2005, seventh edition. Contact: Alternative Link LLC; 1065 Main St., Bldg. C, Las Cruces, NM 88005; phone: (505) 527-0636; http://www.alternativelink.com; mail@alternativelink.com.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Alternative Link LLC, 6121 Indian School Road NE, Suite 131, Albuquerque, NM 87110; Phone: (505) 875-0001;

Toll Free: (877)621-5465; Fax: (505) 875-0002; e-mail: ail@alternativelink.com

AOD2000 - Alcohol and Other Drug Thesaurus: A Guide to Concepts and Terminology in Substance Abuse and Addiction. 3rd. ed. [4 Volumes.] Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism (NIAAA) and Center for Substance Abuse Prevention (CSAP), 2000

Contact: Nancy Winstanley, NIAAA Library, c/o CSR Incorporated 2107 Wilson Blvd., Suite 1000, Arlington, VA 22201; phone: 703-741-7147; e-mail: nwinstanley@csrincorporated.com

---

BI98 - Beth Israel OMR Clinical Problem List Vocabulary. Version 1.0. Boston (MA): Beth Israel Deaconess Medical Center, 1999. Contact: Howard Goldberg, MD.; http://clinquery.bidmc.harvard.edu.

CATEGORY 2 RESTRICTIONS APPLY

Contact: Daniel Z. Sands, M.D., M.P.H., Clinical Systems Integration Architect, Center for Clinical Computing, Beth Israel Deaconess Medical Center, Harvard University, 330 Brookline Avenue, Boston, MA 02215; Phone: 617-667-1510; Fax: 810-592-0716; e-mail: dsands@bidmc.Harvard.edu

---

CCPSS99 - Canonical Clincial Problem Statement System, Version 1.0 June 23, 1999. Contact: sbrown@vumclib.mc.vanderbilt.edu

CATEGORY 3 RESTRICTIONS APPLY

Permission will be freely given for any uses and applications containing CCPSS which are not for sale - i.e. those used internally or given to others without charge.

Contact: Steven Brown, M.D., Associate Professor, Biomedical Informatics, Room 442, Eskind Biomedical Library, Vanderbilt University Medical Center, 2209 Garland Ave Nashville TN 37232-8340; Office: (615) 321-6335; email: sbrown@vumclib.mc.vanderbilt.edu

---

CCS2005 (updated) - Clinical Classifications Software (CCS). Agency for Healthcare Research and Quality (AHRQ), Rockville, MD. Release Date: April 2005. URL: http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp Phone: 301-594-1364.

Contact: Anne Elixhauser, Ph.D., Senior Research Scientist, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850; phone: (301) 427-1411; fax: (301) 594-1430; phone: 1-800-358-9295; email: AElixhau@AHRQ.gov

---

CDT5 - Current Dental Terminology 2005 (CDT 2005). Chicago, IL: American Dental Association, 2005

CATEGORY 3 RESTRICTIONS APPLY

For CDT the following special notice must be displayed:
"For CDT only, copyright 2002 American Dental Association, all rights reserved."

Contact: Ann Pollack, American Dental Association, 211 East Chicago Ave., Chicago, IL 60611-2678; phone: (312)440-2500, ext. 2854

COSTAR_89-95 - Computer-Stored Ambulatory Records (COSTAR). Boston (MA): Massachusetts General Hospital, 1989-1995.

The UMLS Metathesaurus includes terms that were used frequently at 3 COSTAR sites in the years indicated and supplied to NLM by Massachusetts General Hospital.

Contact: G.Octo Barnett, M.D., Laboratory of Computer Science Massachusetts General Hospital, 50 Staniford Street, 5th Floor, Boston, MA 02114; phone: (617) 726-3939; fax: (617) 726-8481; e-mail: Barnett.Octo@mgh.harvard.edu

---

CPM2003 - Medical Entities Dictionary (CPM), Columbia Presbyterian Medical Center Medical Entities Dictionary. New York (NY): Columbia Presbyterian Medical Center, 2003

CATEGORY 2 RESTRICTIONS APPLY

The UMLS Metathesaurus includes a relatively small number of terms created at Columbia Presbyterian Medical Center for the MED, which also includes terms obtained from the UMLS Metathesaurus and other sources.

Contact: James J Cimino M.D., Professor, Department of Medical Informatics, Columbia University, Vanderbilt Clinic VC-5, 622 W. 168th Street, New York NY 10032; phone: (212) 305-8127; fax: (212) 305-3302; e-mail: ciminoj@dbmi.columbia.edu

---

CPT2005 - Current Procedural Terminology (CPT), 4th ed. Chicago (IL): American Medical Association, 2004. http://www.ama-assn.org

CATEGORY 3 RESTRICTIONS APPLY

The following special notic must be displayed:

"CPT™ only Copyright 2004 American Medical Association. All rights Reserved. No fee schedules, basic unit, relative values or related listings are included in CPT™. AMA does not directly or indirectly practice medicine or dispense medical services. AMA assumes no liability for data contained herein.

U.S. Government Rights

This product includes CPT™ which is commercial technical data and/or computer data bases and/or commercial computer software and/or commercial computer software documentation, as applicable which were developed exclusively at private expense by the American Medical Association, 515 North State Street, Chicago, Illinois, 60610. U.S. Government rights to use, modify, reproduce, release, perform, display, or disclose these technical data and/or computer data bases and/or computer software and/or computer software documentation are subject to the limited rights restrictions of DFARS 252.227-7015(b)(2) (June 1995) and/or subject to the restrictions of DFARS 227.7202-1(a) (June 1995) and DFARS 227.7202-3(a) (June 1995), as applicable for U.S. Department of Defense procurements and the limited rights restrictions of FAR 52.227-14 (June 1987) and/or subject to the restricted rights provisions of FAR 52.227-14 (June 1987) and FAR 52.227-19 (June 1987), as applicable, and any applicable agency FAR Supplements, for non-Department of Defense Federal procurements."

Contact: Dorith Brown, CPT Intellectual Property Services, American Medical Association, 515 N. State Street, Chicago, IL 60610; fax: (312)464-5762

CPT01SP - Current Procedural Terminology (CPT), Spanish Translation. 4th ed. Chicago (IL): American Medical Association, 2000. http://www.ama-assn.org.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Dorith Brown, CPT Intellectual Property Services, American Medical Association, 515 N. State Street, Chicago IL 60610; Fax: (312) 464-5762

CSP2005 (updated) - Computer Retrieval of Information on Scientific Projects (CRISP). Bethesda (MD): National Institutes of Health, Division of Research Grants, Research Documentation Section, 2005.

Contact: Dorrette Finch, Director, Division of Research Documentation, ORA, OER, National Institutes of Health, 6701 Rockledge Drive, Bethesda MD 20892-7983; email: dw33v@nih.gov

---

CST95 - Coding Symbols for Thesaurus of Adverse Reaction Terms (COSTART). 5th ed. Rockville (MD): U.S. Food and Drug Administration, Center for Drug Evaluation and Research, 1995.

COSTART has been superseded by the Medical Dictionary for Regulatory Activities (MedDRA) Terminology.

Contact: National Technical Information Service. http://www.ntis.gov/fcpc/cpn5580.htm

---

CTCAEV3 (new) - Common Terminology Criteria for Adverse Events v3.0, Cancer Therapy Evaluation Program (CTEP), National Cancer Institute. Bethesda, MD: National Cancer Institute, Dec. 12, 2003.

Contact: M. Haber, National Cancer Institute, Building 6116, 6116 Exec Blvd, Room 3124, Rockville, MD; Phone: 301-594-9185; Fax: 301-480-8105; Email: mhaber@mail.nih.gov

---

DDB00 - Diseases Database 2000. May, 2000. London (England): Medical Object Oriented Software Enterprises Ltd., 2000. Contact: Malcolm Duncan <mhduncan@compuserve.com>; http://www.diseasesdatabase.com/.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Malcolm H. Duncan, Medical Object Oriented Software Enterprises Ltd, Unit 36c Marryat Square, Fulham, London SW6 6UA, UK; Tel: 44 (0)20 7381 4220; Mobile: 07710 483088; email: mhduncan@compuserve.com; URL: http://www.diseasesdatabase.com/

DMDICD10_1995 - Internationale Klassifikation der Krankheiten 10 [German translation of ICD10]. Germany: Deutsches

Institut fuer Medizinische Dokumentation und Information, 1998.

Contact: Dr. Michael Schopen; e-mail: schopen@dimdi.de; Deutsches Institut fur Medizinische Dokumentation und Information (DIMDI), Postfach 420580, D-50899 Koln, Germany; phone: 49-221-472-4252; fax: 49-221-41-1429

DMDUMD_1996 - Die Nomenklatur fuer Medizinprodukte UMDNS [German translation of UMDNS]. Germany: Deutsches Institut fuer Medizinische Dokumentation und Information, 1996.

Contact: Dr. Michael Schopen; Deutsches Institut fur Medizinische Dokumentation und Information (DIMDI), Postfach 420580, D-50899 Koln, Germany; tel: 49-221-472-4252; FAX: 49-221-41-1429; e-mail: helpdesk@dimdi.de; URL: www. dimdi.de

---

DSM3R_1987 - Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R). 3rd ed. rev. Washington (DC): American Psychiatric Association, 1987.

Contact: Customer Service, American Psychiatric Press Inc.,1400 K Street N.W., Washington DC 20005; e-mail: csdept@appi.org

---

DSM4_1994 - Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). 4th ed. Washington (DC): American Psychiatric Association, 1994.

The APA usually charges small administrative fees for copyright permissions, but these may be waived for research purposes. All users should apply for permission in writing or by email to:

Contact: Customer Service, American Psychiatric Press Inc., 1400 K Street N.W., Washington DC 20005; e-mail: csdept@appi.ord

---

DXP94 - DXplain (An expert diagnosis program). Boston (MA): Massachusetts General Hospital, 1994.

Contact: G. Octo Barnett, M.D., Laboratory of Computer Science, Massachusetts General Hospital, 50 Staniford Street, 5th Floor, Boston, MA 02114; phone: (617) 726-3939; fax: (617) 726-8481; e-mail: Barnett.Octo@mgh.Harvard

---

GO2004_12_20 (updated) - Gene Ontology. The Gene Ontology Consortium, March 2, 2004. Available from: http://www. geneontology.org/#cite_go.

Contact: http://www.geneontology.org/#cite_go; go@geneontology.org

---

HCDT5 - HCPCS Version of Current Dental Terminology 2005 (CDT-2005). Baltimore, MD: U.S. Centers for Medicare & Medicaid, 2005.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Ann Pollack, American Dental Association; phone: (312)440-2500, ext. 2854

HCPCS05 - Healthcare Common Procedure Coding System (HCPCS). Baltimore, MD: U.S. Centers for Medicare & Medicaid Services, 2005.

The American Medical Association's CPT™ codes in HCPCS have a Source Abbreviation of HCPT04. The American Dental Association's CDT codes in HCPCS have a Source Abbreviation of HCDT4.

Contact: Cynthia Hake, Centers for Medicare & Medicaid Services (CMS) 7500 Security Blvd., Mailstop C5-09-16, Baltimore MD 21244; e-mail:CHake@cms.hhs.gov; phone: (410) 786-3404

HCPT05 - Version of Physicians' Current Procedural Terminology (CPT) included in the Healthcare Common Procedure Coding System (HCPCS), 2005.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Cynthia Hake, Centers for Medicare & Medicaid Services (CMS) 7500 Security Blvd., Mailstop C5-09-16, Baltimore MD 21244; e-mail:CHake@cms.hhs.gov; phone: (410) 786-3404

HHC2003 - Saba, Virginia. Home Health Care Classification of Nursing Diagnoses and Interventions. Washington (DC): Georgetown University, 2003.

CATEGORY 1 RESTRICTIONS APPLY

Contact: Virginia K. Saba, EdD, RN, FAAN, FACMI, LL, Distinguished Scholar, Adjunct, Georgetown University, Washington, DC; Professor, Adjunct, USUHS, Bethedsa, MD 2332 South Queen Street, Arlington VA 22202; Tel: 703-521-6132; Fax: 703-521-3866; e-mail: vsaba@worldnet.att.net

---

HL7V2.5_2003_08_30 (new) - Health Level Seven Vocabulary (HL7). Ann Arbor (MI): Health Level Seven, 1998-2002. Contact: Mark McDougall, Executive Director, Health Level Seven; 3300 Washtenaw Avenue, Suite 227, Ann Arbor, MI 48104-4250; Phone: (734)677-7777; Fax: (734)677-6622; Email: HQ@HL7.ORG ; Web site: www.HL7.ORG.

Contact: Health Level Seven, 3300 Washtenaw Avenue, Suite 227, Ann Arbor MI 48104-4250; phone: (734)677-7777; fax: (734)677-6622; e-mail: HQ@HL7.ORG

HL7V3.0_2004_12_09 (new) - Health Level Seven Vocabulary (HL7). Ann Arbor (MI): Health Level Seven, 1998-2002. Contact: Mark McDougall, Executive Director, Health Level Seven; 3300 Washtenaw Avenue, Suite 227, Ann Arbor, MI 48104-4250; Phone: (734)677-7777; Fax: (734)677-6622; Email: HQ@HL7.ORG ; Web site: www.HL7.ORG.

Contact: Health Level Seven, 3300 Washtenaw Avenue, Suite 227, Ann Arbor MI 48104-4250; phone: (734)677-7777; fax: (734)677-6622; e-mail: HQ@HL7.ORG

HLREL_1998 - ICPC2E-ICD10 relationships from Dr. Henk Lamberts (HLREL), 1998. University of Amsterdam. Contact: H.Lamberts@AMC.UVA.NL.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Henk Lamberts (HLREL), University of Amsterdam; email: H.Lamberts@AMC.UVA.NL.

HUGO_2005_04 (updated) - HUGO Gene Nomenclature, HUGO Gene Nomenclature Committee, Department of Biology, University College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, UK. Tel: 44-20-7679-5027 Fax: 44-20-7387-3496 e-mail: nome@galton.ucl.ac.uk

Contact: nome@galton.ucl.ac.uk

ICD10_1998 - International Statistical Classification of Diseases and Related Health Problems (ICD-10). 10th rev. Geneva (Switzerland): World Health Organization, 1998.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Office of Publications, World Health Organization, 1211 Geneva 27, Switzerland

---

ICD10AE_1998 - International Statistical Classification of Diseases and Related Health Problems (ICD-10): Americanized Version. 10th rev. Geneva (Switzerland): World Health Organization, 1998.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Office of Publications, World Health Organization, 1211 Geneva 27, Switzerland

---

ICD10AM_2000 - International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian Modification; 2nd Edition, published January 2000. Developed and Maintained by the National Centre for Classification in Health, University of Sydney, Faculty of Health Sciences. PO Box 170 Lidcombe, NSW, Australia 1825. Phone: +61 2 9351 9461. http://www.cchs.usy.edu.au/ncch/

CATEGORY 3 RESTRICTIONS APPLY

Contact: National Centre for Classification in Health, University of Sydney, Faculty of Health Sciences; PO Box 170 Lidcombe, NSW, Australia 1825; Phone: +61 2 9351 9461

ICD10AMAE_2000 - International Statistical Classification of Diseases and Related Health Problems, Australian Modification (ICD-10-AM), Americanized English Equivalents, produced by NLM. Bethesda (MD): National Library of Medicine, UMLS project, 2000

CATEGORY 3 RESTRICTIONS APPLY

Contact: National Centre for Classification in Health University of Sydney Faculty of Health Sciences PO Box 170 Lidcombe, NSW Australia 1825; phone: +61 2 9351 9461

ICD10DUT_200403 - Hirs, W., H.W. Becker, C. van Boven, S.K. Oskam, I.M. Okkes, H. Lamberts. ICD-10, Dutch Translation, 200403. Amsterdam: Department of General Practice, Academic Medical Center/University of Amsterdam, Dutch College of General Practitioners (NHG), March 2004.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Inge M. Okkes, PhD, Department of Family Practice, Division Clinical Methods & Public Health, Academic Medical Center/University of Amsterdam, Meibergdreef 15 (Room J2-213), 1105 AZ AMSTERDAM, The Netherlands; Tel: 31 20 566 4660; E-mail: i.m.okkes@amc.uva.nl

---

ICD9CM_2006 (updated) - ICD-9-CM [computer file]: international classification of diseases, ninth revision, clinical modification. Baltimore, MD: U.S. Department of Health and Human Services, Centers for Medicare & Medicaid Services, effective October 1, 2005.

NLM has generated fully specified titles for ICD-9-CM codes in cases in which the official ICD- 9-CM titles consist of extensions to higher levels in the ICD-9-CM hierarchy. The fully specified names were produced with reasonable care, but have not yet been reviewed and approved by the producers of ICD-9-CM.

Contact: Contact for Diseases: Donna Pickett, National Center for Health Statistics; e-mail: dfp4@cdc.gov; Contact for Procedures: Patricia Brooks, Health Care Financing Administration; e-mail: pbrooks@hcfa.gov;

ICD9CM_2005 - ICD-9-CM [computer file]: international classification of diseases, ninth revision, clinical modification. Version 22. Baltimore, MD: U.S. Department of Health and Human Services, Centers for Medicare & Medicaid Services, effective October 1, 2004.

NLM has generated fully specified titles for ICD-9-CM codes in cases in which the official ICD- 9-CM titles consist of extensions to higher levels in the ICD-9-CM hierarchy. The fully specified names were produced with reasonable care, but have not yet been reviewed and approved by the producers of ICD-9-CM.

Contact: Contact for Diseases: Donna Pickett, National Center for Health Statistics; e-mail: dfp4@cdc.gov; Contact for Procedures: Patricia Brooks, Health Care Financing Administration; e-mail: pbrooks@hcfa.gov;

ICPC93 - The International Classification of Primary Care (ICPC). Denmark: World Organisation of Family Doctors, 1993.

This year, the Metathesaurus has also included translations of ICPC93 in the following languages:

- Basque (ICPCBAQ_1993),
- Danish (ICPCDAN_1993),
- Dutch (ICPCDUT_1993),
- Finnish (ICPCFIN_1993),
- French (ICPCFRE_1993),
- German (ICPCGER_1993),
- Hebrew (ICPCHEB_1993),
- Hungarian (ICPCHUN_1993),
- Italian (ICPCITA_1993),
- Norwegian (ICPCNOR_1993),
- Portuguese (ICPCPOR_1993),
- Spanish (ICPCSPA_1993), and
- Swedish (ICPCSWE_1993).

Contact: Henk Lamberts; e-mail: H.Lamberts@AMC.UVA.NL or Inge Hofmans-Okkes; e-mail: I.m.okkes@amc.una.nl; University of Amsterdam

---

ICPC2EDUT_200203 - Hirs, W., H.W. Becker, C. van Boven, S.K. Oskam, I.M. Okkes, H. Lamberts. International Classification of Primary Care 2E: 2nd ed. electronic. Dutch Translation. Amsterdam: Department of General Practice, Academic Medical Center/University of Amsterdam, Dutch College of General Practitioners (NHG), March 2002

CATEGORY 3 RESTRICTIONS APPLY

Contact: Prof. Dr. H.Lamberts and Inge M. Okkes, PhD, Department of Family Practice, Division Clinical Methods & Public Health, Academic Medical Center/University of Amsterdam, Meibergdreef 15 (Room J2-213), 1105 AZ Amsterdam, The Netherlands; Tel: 31 20 566 4711

ICPC2EENG_200203 - International Classification of Primary Care / prepared by the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians (WONCA), known more briefly as the World Organization of Family Doctors. 2nd ed. Henk Lamberts and Inge Hofmans-Okkes, 2002

CATEGORY 3 RESTRICTIONS APPLY

Contact: Prof. Dr. H.Lamberts and Inge M. Okkes, PhD, Department of Family Practice, Division Clinical Methods & Public Health, Academic Medical Center - University of Amsterdam, Meibergdreef 15 (Room J2-213), 1105 AZ Amsterdam, The Netherlands; Tel: 31 20 566 4711

ICPC2ICD10DUT_200412 (updated) - International Classification of Primary Care / prepared by the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/ Family Physicians (WONCA), known more briefly as the World Organization of Family Doctors. 2nd ed. Henk Lamberts and Inge Hofmans-Okkes, Dutch, 2002

CATEGORY 3 RESTRICTIONS APPLY

Contact: Prof. Dr. H. Lamberts and Inge M. Okkes, PhD, Department of Family Practice, Division of Clinical Methods & Public Health, Academic Medical Center - University of Amsterdam, Meibergdreef 15 (Room J2-213), 1105 AZ Amsterdam, The Netherlands; Tel: 31 20 566 4711

ICPC2ICD10ENG_200412 (updated) - International Classification of Primary Care / prepared by the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/ Family Physicians (WONCA), known more briefly as the World Organization of Family Doctors. 2nd ed. Henk Lamberts and Inge Hofmans-Okkes, 2002

CATEGORY 3 RESTRICTIONS APPLY

Contact: Prof. Dr. H. Lamberts and Inge M. Okkes, PhD, Department of Family Practice, Division of Clinical Methods & Public Health, Academic Medical Center/University of Amsterdam, Meibergdreef 15 (Room J2-213), 1105 AZ Amsterdam, The Netherlands; Tel: 31 20 566 4711

ICPC2P_2000 - International Classification of Primary Care, Version 2-Plus, Australian Modification. January, 2000

CATEGORY 3 RESTRICTIONS APPLY

Contact: HelenaB@genprac.wsahs.nsw.gov.au

ICPCBAQ_1993 - The International Classification of Primary Care (ICPC). Basque Translation. Denmark: World Organisation of Family Doctors, 1993.

Contact: Henk Lamberts; e-mail: H.Lamberts@AMC.UVA.NL or Inge Hofmans-Okkes; e-mail: I.m.okkes@amc.una.nl; University of Amsterdam.

ICPCDAN_1993 - The International Classification of Primary Care (ICPC). Danish Translation. Denmark: World Organisation of Family Doctors, 1993.

Contact: Henk Lamberts; e-mail: H.Lamberts@AMC.UVA.NL or Inge Hofmans-Okkes; e-mail: I.m.okkes@amc.una.nl; University of Amsterdam

---

ICPCDUT_1993 - The International Classification of Primary Care (ICPC). Dutch Translation. Denmark: World Organisation of Family Doctors, 1993.

Contact: Henk Lamberts; e-mail: H.Lamberts@AMC.UVA.NL or Inge Hofmans-Okkes; e-mail: I.m.okkes@amc.una.nl; University of Amsterdam.

ICPCFIN_1993 - The International Classification of Primary Care (ICPC). Finnish Translation. Denmark: World

Organisation of Family Doctors, 1993.

Contact: Henk Lamberts; e-mail: H.Lamberts@AMC.UVA.NL or Inge Hofmans-Okkes; e-mail: I.m.okkes@amc.una.nl; University of Amsterdam.

ICPCFRE_1993 - The International Classification of Primary Care (ICPC). French Translation. Denmark: World Organisation of Family Doctors, 1993.

Contact: Henk Lamberts; e-mail: H.Lamberts@AMC.UVA.NL or Inge Hofmans-Okkes; e-mail: I.m.okkes@amc.una.nl; University of Amsterdam.

ICPCGER_1993 - The International Classification of Primary Care (ICPC). German Translation. Denmark: World Organisation of Family Doctors, 1993.

Contact: Henk Lamberts; e-mail: H.Lamberts@AMC.UVA.NL or Inge Hofmans-Okkes; e-mail: I.m.okkes@amc.una.nl; University of Amsterdam.

ICPCHEB_1993 - The International Classification of Primary Care (ICPC). Hebrew Translation, Denmark: World Organisation of Family Doctors, 1993

Contact: Henk Lamberts; e-mail: H.Lamberts@AMC.UVA.NL or Inge Hofmans-Okkes; e-mail: I.m.okkes@amc.una.nl; University of Amsterdam.

ICPCHUN_1993 - The International Classification of Primary Care (ICPC). Hungarian Translation. Denmark: World Organisation of Family Doctors, 1993.

Contact: Henk Lamberts; e-mail: H.Lamberts@AMC.UVA.NL or Inge Hofmans-Okkes; e-mail: I.m.okkes@amc.una.nl; University of Amsterdam.

ICPCITA_1993 - The International Classification of Primary Care (ICPC). Italian Translation. Denmark: World Organisation of Family Doctors, 1993.

Contact: Henk Lamberts; e-mail: H.Lamberts@AMC.UVA.NL or Inge Hofmans-Okkes; e-mail: I.m.okkes@amc.una.nl; University of Amsterdam.

ICPCNOR_1993 - The International Classification of Primary Care (ICPC). Norwegian Translation. Denmark: World Organisation of Family Doctors, 1993.

Contact: Henk Lamberts; e-mail: H.Lamberts@AMC.UVA.NL or Inge Hofmans-Okkes; e-mail: I.m.okkes@amc.una.nl; University of Amsterdam.

ICPCPAE_2000 - International Classification of Primary Care, Version 2-Plus, Australian Modification. Americanized English Equivalents, January, 2000. Produced by NLM. Bethesda (MD): National Library of Medicine, UMLS project

CATEGORY 3 RESTRICTIONS APPLY

Contact: HelenaB@genprac.wsahs.nsw.gov.au

ICPCPOR_1993 - The International Classification of Primary Care (ICPC). Portuguese Translation. Denmark: World Organisation of Family Doctors, 1993.

Contact: Henk Lamberts; e-mail: H.Lamberts@AMC.UVA.NL or Inge Hofmans-Okkes; e-mail: I.m.okkes@amc.una.nl; University of Amsterdam.

ICPCSPA_1993 - The International Classification of Primary Care (ICPC). Spanish Translation. Denmark: World Organisation of Family Doctors, 1993.

Contact: Henk Lamberts; e-mail: H.Lamberts@AMC.UVA.NL or Inge Hofmans-Okkes; e-mail: I.m.okkes@amc.una.nl; University of Amsterdam.

ICPCSWE_1993 - The International Classification of Primary Care (ICPC). Swedish Translation. Denmark: World Organisation of Family Doctors, 1993.

Contact: Henk Lamberts; e-mail: H.Lamberts@AMC.UVA.NL or Inge Hofmans-Okkes; e-mail: I.m.okkes@amc.una.nl; University of Amsterdam.

JABL99 - Online Congenital Multiple Anomaly/Mental Retardation Syndromes, 1999.

CATEGORY 1 RESTRICTIONS APPLY

Contact: Stanley Jablonski; National Library of Medicine, 8600 Rockville Pike, Bethesda MD 20894; email: stanley_jablonski@nlm.nih.gov

---

LCH90 - Library of Congress Subject Headings. 12th ed. Washington (DC): Library of Congress, 1989.

There are later editions of this source that are not reflected in the UMLS Metathesaurus. This source has considerable non-biomedical content and will never be included in the Metathesaurus in its entirety.

Contact: http://www.lcweb.loc.gov

LNC215 (updated) - Logical Observation Identifier Names and Codes (LOINC). Version 2.15. Indianapolis (IN): The Regenstrief Institute, June 6, 2005

Contact: kmercer@regenstrief.org

MBD05 - MEDLINE Backfiles (1995-1999). Bethesda (MD): National Library of Medicine. Contact: http://www.nlm.nih.gov.

Contact: National Library of Medicine. Bethesda, MD; http://www.nlm.nih.gov

---

MCM92 - Glossary of Methodologic Terms for Clinical Epidemiologic Studies of Human Disorders. Canada: McMaster University, 1992.

Contact: R. Brian Haynes, M.D., Ph.D.; e-mail: bhaynes@mcmaster.ca; Clinical Epidemiology & Biostatistics and Medicine, Faculty of Health Sciences, McMaster University, Room 2C10B, 1200 Main Street, West Hamilton Ontario, Canada L8N 3Z5; phone (905) 525-9140

MDDB_2003_03 - Master Drug Data Base, 2003

CATEGORY 3 RESTRICTIONS APPLY

Contact: Karen Eckert, RPh, Medi-Span, 8425 Woodfield Crossing Blvd., Suite 490, Indianapolis IN 46240; phone (800) 388-8884; fax (317) 735-5390; email: Keckert@drugfacts.com; ms-support@drugfacts.com; http://www.medi-span.com/products/product_mddb.asp

---

MDR80 (updated) - MedDRA [electronic resource] : medical dictionary for regulatory activities terminology. Version 8.0. Reston, VA : Northrop Grumman, MedDRA MSSO, September 2004.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Elizabeth d'Alelio; Customer Operations Manager; MedDRA MSSO; 12011 Sunset Hills Rd. 8A16; Reston, VA 20910; Elizabeth.d'Alelio@trw.com; phone: (703)345-8821; fax: (703)345-7791

MDR60 - Medical Dictionary for Regulatory Activities Terminology (MedDRA) Version 6.0, March, 2003. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). http://meddramsso.com

CATEGORY 3 RESTRICTIONS APPLY

Contact: Elizabeth d'Alelio; Customer Operations Manager; MedDRA MSSO; 12011 Sunset Hills Rd. 8A16; Reston, VA 20910; Elizabeth.d'Alelio@trw.com; phone: (703)345-8821; fax: (703)345-7791

MDRAE80 (updated) - MedDRA : medical dictionary for regulatory activities terminology. American English Equivalents, Version 8.0. Bethesda, MD: U.S. National Library of Medicine, September 1, 2004.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Elizabeth d'Alelio; Customer Operations Manager; MedDRA MSSO; 12011 Sunset Hills Rd. 8A16; Reston, VA 20910; Elizabeth.d'Alelio@trw.com; phone: (703)345-8821; fax: (703)345-7791

MDRDUT71 - Medical Dictionary for Regulatory Activities Terminology (MedDRA) Version 7.1, Dutch Edition. International conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Reston, VA: MedDRA MSSO, September 2004

CATEGORY 3 RESTRICTIONS APPLY

Contact: Customer Operations Manager; MedDRA MSSO; 12011 Sunset Hills Rd. 8A16; Reston, VA 20910; Elizabeth d'Alelio@trw.com; phone: (703)345-8821; fax: (703)345-7791

MDREA80 (updated) - MedDRA : medical dictionary for regulatory activities terminology, american english equivalents with expanded Abbreviations. Version 8.0. Bethesda, MD: U.S. National Library of Medicine, September 1, 2004.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Elizabeth d'Alelio; Customer Operations Manager; MedDRA MSSO; 12011 Sunset Hills Rd. 8A16; Reston, VA 20910; Elizabeth.d'Alelio@trw.com; phone: (703)345-8821; fax: (703)345-7791

MDREX80 (updated) - MedDRA : medical dictionary for regulatory activities terminology, with expanded abbreviations. Version 8.0. Bethesda, MD: U.S. National Library of Medicine, September 1, 2004.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Elizabeth d'Alelio; Customer Operations Manager; MedDRA MSSO; 12011 Sunset Hills Rd. 8A16; Reston, VA 20910; Elizabeth.d'Alelio@trw.com; phone: (703)345-8821; fax: (703)345-7791

MDRFRE71 - Medical Dictionary for Regulatory Activities Terminology (MedDRA) Version 7.1, French Edition. International conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Reston, VA: MedDRA MSSO, September 2004.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Customer Operations Manager; MedDRA MSSO; 12011 Sunset Hills Rd. 8A16; Reston, VA 20910; Elizabeth.d'Alelio@trw.com; phone: (703)345-8821; fax: (703)345-7791

MDRGER80 (updated) - Medical Dictionary for Regulatory Activities Terminology (MedDRA) Version 8.0, German Edition. International conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Reston, VA: MedDRA MSSO, August 2005.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Kathryn Studeman Customer Operations, MedDRA MSSO, 12011 Sunset Hills Rd., VAR1/7B34/MSSO, Reston, VA 20190-3285; phone: 703-345-8175; fax: 703-345-7791; Email: kathryn.studeman@ngc.com

---

MDRPOR71 - Medical Dictionary for Regulatory Activities Terminology (MedDRA) Version 7.1, Portuguese Edition. International conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Reston, VA: MedDRA MSSO, September 2004.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Customer Operations Manager; MedDRA MSSO; 12011 Sunset Hills Rd. 8A16; Reston, VA 20910; Elizabeth.d'Alelio@trw.com; phone: (703)345-8821; fax: (703)345-7791

MDRSPA71 - Medical Dictionary for Regulatory Activities Terminology (MedDRA) Version 7.1, Spanish Edition. International conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Reston, VA: MedDRA MSSO, September 2004.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Elizabeth d'Alelio; Customer Operations Manager; MedDRA MSSO; 12011 Sunset Hills Rd. 8A16;Reston, VA 20910; Elizabeth.d'Alelio@trw.com; phone: (703)345-8821; fax: (703)345-7791

MED05 - MEDLINE Current Files (2000-2005). Bethesda (MD): National Library of Medicine. Contact: http://www.nlm.nih.gov.

Contact: Jan Willis, National Library of Medicine, 38A-4th floor, 8600 Rockville Pike, Bethesda MD 20894; phone: 301-496-7715; e-mail: jwillis@nlm.nih.gov

---

MEDLINEPLUS_20040814 - MedlinePlus Health Topics. Bethesda (MD): National Library of Medicine, August 14, 2004.

Contact: Naomi Miller; e-mail: millern@mail.nlm.nih.gov

MIM93 - Online Mendelian Inheritance in Man (OMIM). Baltimore (MD): Johns Hopkins University, Center for Biotechnology Information, 1994.

CATEGORY 1 RESTRICTIONS APPLY

To date the UMLS Metathesaurus contains a relatively small amount of data from this source.

Contact: Joanna Amberger, Johns Hopkins University; e-mail: joanna@ncbi.nlm.nih.gov; Leigh Penfield, OMIM, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins Hospital, Blalock 1007, 600 N. Wolfe St., Baltimore MD 21287-4922; phone 410-955-0313; fax 410-955-4999; email: techlicense@jhmi.edu.

MMSL_2004_03 - Medisource Lexicon. Multum Information Services, Inc., Denver, CO. Release Date: March 1, 2004. URL: http://www.multum.com/ Phone: 888-633-4772 x1420.

CATEGORY 1 RESTRICTIONS APPLY

Contact: Multum Information Services, 3200 Cherry Creek South Drive, Suite 300, Denver CO 80209

MMX01 - Micromedex DRUGDEX, 2001. URL: http://www.micromedex.com/ Phone: 800-525-9083.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Micromedex, 6200 South Syracuse Way, Suite 300, Englewood CO 80111-4740; phone: (800) 525-9083; e-mail: info@mdx.com

---

MSH2006_2005_08_01 (updated) - Medical Subject Headings (MeSH). Bethesda (MD): National Library of Medicine, 2006

This source has been translated into many languages. To date, eight of the translations have been incorporated into the UMLS Metathesaurus.

Contact: Stuart Nelson, M.D., Head, MeSH Section; e-mail: nelson@nlm.nih.gov

---

MSH2005_2005_01_17 (updated) - Medical Subject Headings (MeSH). Bethesda (MD): National Library of Medicine, 2005

This source has been translated into many languages. To date, eight of the translations have been incorporated into the UMLS Metathesaurus.

Contact: Stuart Nelson, M.D., Head, MeSH Section; e-mail: nelson@nlm.nih.gov

---

MSHCZE2004 (updated) - Czech translation of Medical Subject Headings (MeSH) 2004. Prague: Dept. of Bibliography, National Library of Medicine, 2004

CATEGORY 3 RESTRICTIONS APPLY

Contact: Marie Votipkova, Dept. of Bibliography, National Library of Medicine, Prague, Sokolska 54, Czech Republic; e-mail: votipkov@nlk.cz; URL: www.nlk.cz

---

MSHDUT2005 (updated) - Nederlandse vertaling van MeSH [Dutch translation of MeSH), 2005. Amsterdam: Nederlands Tijdschrift voor Geneeskunde [Dutch Journal of Medicine], 2005.

CATEGORY 3 RESTRICTIONS APPLY

Contact: A.J.P.M.Overbeke, overbeke@ntvg.nl, * 20 662 0150

MSHFIN2005 (updated) - Finnish translations of Medical Subject Headings (MeSH), 2005. Helsinki: Finnish Medical Society Duodecim, 2005.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Peter Nyberg, MD; e-mail: peter.nyberg@sll.fimnet.fi

---

MSHFRE2005 - Thesaurus Biomedical Francais/Anglais [French translation of MeSH]. Paris (France): Institut National de la Sante et Recherche Medicale, 2005

CATEGORY 3 RESTRICTIONS APPLY

Contact: Claudie Hasenfuss, Institut National de la Sante et Recherche Medicale; phone: 33-1-45-59-52-91; e-mail: hasenfus@vjf.inserm.fr; http://www.inserm.fr

---

MSHGER2005 - German translation of Medical Subject Headings (MeSH), 2005. Cologne: Deutsches Institut fur Medizinische Dokumentation und Information, 2005.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Dr. Michael Schopen, Deutsches Institut fur Medizinische Dokumentation und Information; phone: 49-221-472-4325; e-mail: schopen@dimdi.de

---

MSHITA2005 - Italian translation of Medical Subject Headings (MeSH), 2005. Rome: Istituto Superiore di Sanita, Settore Documentazione, 2005.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Dr. Adriana Dracos, Viale Regina Elena, 229 00616 Rome Italy; phone:39-06-49902280; fax: 39-06-49387117; e-mail: dracos@iss.it

---

MSHJPN2005 (updated) - JAMAS Japanese Medical Thesaurus (JJMT). Tokyo: Japan Medical Abstracts Society; Igaku-Chuo-Zasshi, 2005.

Contact: Tsuneki Shinohara, MD, President, Japan Medical Abstracts Society, 2-5-18, Tokaido-Higashi, Suginami, Tokyo 168-0072, Japan; e-mail: shinohara@jamas.gr.jp

---

MSHPOR2005 - Descritores em Ciencias da Saude [Portuguese translation of Medical Subject Headings (MeSH)], 2004. Sao Paulo (Brazil): Latin American and Caribbean Center on Health Sciences Information. BIREME/PAHO/WHO, 2005.

Contact: Elenice de Castro; e-mail: elenice@brm.bireme.br

MSHRUS2005 (updated) - Russian Translation of Medical Subject Headings (MeSH). Moscow: State Central Scientific Medical Library, 2005

Contact: Natalya Lomova, Head of Department State Central Scientific Medical Library, Nakhimovsky prospekt 49, Moscow, GSP 117418, Russia; e-mail:vld@vyugin.mccme.rssi.ru

---

MSHSPA2005 - Descriptores en Ciencias de la Salud [Spanish translation of Medical Subject Headings (MeSH)], 2004. Sao Paulo: Latin American and Caribbean Center on Health Sciences Information. BIREME/PAHO/WHO, 2005.

Contact: Elenice de Castro; e-mail: elenice@brm.bireme.br

MSHSWE2004 - Swedish translation of Medical Subject Headings (MeSH), 2004. Stockholm: Karolinska Institutet, 2004.

Contact: Ylva Gavel, Library of Karolinska Institutet, Box 200, SE-171 77 Stockholm, Sweden; phone: +46 (0)8-524 84 000, fax: +46 (0)8-524 84 310; e-mail: Ylva.Gavel@kib.ki.se

---

MTH - UMLS Metathesaurus. Bethesda, MD: National Library of Medicine.

Concept names with this source abbreviation were created by NLM to facilitate creation of the UMLS Metathesaurus. There are relatively few of them.

Contact: Jan Willis, National Library of Medicine, UMLS Support, 38A-4th floor, 8600 Rockville Pike, Bethesda MD 20894; phone: 301-496-7715; e-mail: jwillis@nlm.nih.gov

---

MTHCH05 - Metathesaurus Hierarchical CPT Terms (these terms were created by the NLM to provide contextual information for CPT). Bethesda (MD): National Library of Medicine, 2004

Contact: Jan Willis, National Library of Medicine, 38A-4th floor, 8600 Rockville Pike, Bethesda MD 20894; phone: 301-496-7715; e-mail: jwillis@nlm.nih.gov

---

MTHFDA_2005_03_31 (updated) - Metathesaurus Forms of FDA National Drug Code Directory, 2005_03_31. Bethesda, MD: National Library of Medicine, 2005.

Concept names with this source abbreviation were created by NLM to provide contextual information for FDA NDC terms.

Contact: Jan Willis, National Library of Medicine, UMLS Support, 38A-4th fl, 8600 Rockville Pike, Bethesda MD 20894; phone: 301-496-7715; e-mail: jwillis@nlm.nih.gov. NOTE: Users must also obtain rights to use the parent source.

MTHHH05 - Metathesaurus Hierarchical HCPCS Terms. Bethesda (MD): National Library of Medicine, 2005

Concept names with this source abbreviation were created by NLM to provide contextual information for HCPCS.

Contact: Jan Willis, U.S. National Library of Medicine, UMLS Support/MEDLARS Management Section, 8600 Rockville Pike, Bethesda MD 20894; email: jwillis@nlm.nih.gov

---

MTHHL7V2.5_2003_08 (new) - HL7 Vocabulary Version 2.5, 7-bit equivalents created by the National Library of Medicine. Bethesda (MD): August 30, 2003

Contact: Jan Willis, National Library of Medicine, UMLS Support, 38A-4th floor, 8600 Rockville Pike, Bethesda MD 20894; phone: 301-496-7715; email: jwillis@nlm.nih.gov

---

MTHHL7V3.0_2004_12 (new) - HL7 Vocabulary Version 3.0, 7-bit equivalents created by the National Library of Medicine. Bethesda (MD): December 9, 2004.

Contact: Jan Willis, National Library of Medicine, UMLS Support, 38A-4th floor, 8600 Rockville Pike, Bethesda MD 20894; phone: 301-496-7715; email: jwillis@nlm.nih.gov

---

MTHICD9_2006 (updated) - Metathesaurus additional entry terms for ICD-9-CM [computer file]: international classification of diseases, ninth revision, clinical modification. Bethesda, MD: U.S. Dept. of Health and Human Services, Public Health Service, National Institutes of Health, National Library of Medicine, September 2005.

Contact: Jan Willis, National Library of Medicine, 38A-4th floor, 8600 Rockville Pike, Bethesda MD 20894; phone: 301-496-7715; e-mail: jwillis@nlm.nih.gov

---

MTHICPC2EAE_200203 - Henk Lamberts and Inge Hofmans-Okkes. International Classification of Primary Care 2nd Edition, Electronic, 2E, American English Equivalents. Amsterdam: International Classification of Primary Care / prepared by the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians (WONCA), known more briefly as the World Organization of Family Doctors, 2002

CATEGORY 3 RESTRICTIONS APPLY

Concept names with this source abbreviation were created by NLM to provide contextual information for ICPC2E terms.

Contact: Prof. Dr. H.Lamberts and Inge M. Okkes, PhD, Department of Family Practice, Division Clinical Methods & Public Health, Academic Medical Center/University of Amsterdam, Meibergdreef 15 (Room J2-213), 1105 AZ Amsterdam, The Netherlands; Tel: 31 20 566 4711

MTHICPC2ICD107B_0412 (updated) - International Classification of Primary Care / prepared by the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/ Family Physicians (WONCA), known more briefly as the World Organization of Family Doctors. 2nd ed. Henk Lamberts and Inge Hofmans-Okkes, 7-bit Equivalents, 2002

CATEGORY 3 RESTRICTIONS APPLY

Contact: Prof. Dr. H. Lamberts and Inge M. Okkes, PhD, Department of Family Practice, Division of Clinical Methods & Public Health, Academic Medical Center/University of Amsterdam, Meibergdreef 15 (Room J2-213), 1105 AZ Amsterdam, The Netherlands; Tel: 31 20 566 4711

MTHICPC2ICD10AE_0412 (updated) - International Classification of Primary Care / prepared by the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/ Family Physicians (WONCA), known more briefly as the World Organization of Family Doctors. 2nd ed. Henk Lamberts and Inge Hofmans-Okkes, American English Equivalents, 2002

CATEGORY 3 RESTRICTIONS APPLY

Contact: Prof. Dr. H. Lamberts and Inge M. Okkes, PhD, Department of Family Practice, Division of Clinical Methods & Public Health, Academic Medical Center/University of Amsterdam, Meibergdreef 15 (Room J2-213), 1105 AZ Amsterdam, The Netherlands; Tel: 31 20 566 4711

MTHMDRSPA71 - Methathesaurus Forms of Medical Dictionary for Regulatory Activities Terminology (MedDRA) Version 7.1, Spanish Edition. Bethesda, MD: National Library of Medicine, September 2004.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Jan Willis, National Library of Medicine, UMLS Support, 38A-Room 4S-2, 8600 Rockville Pike, Bethesda MD 20894; phone 301-496-7715; email: jwillis@nlm.nih.gov. NOTE: Users must also obtain the rights to use the parent source.

MTHMSHSPA2005 (new) - Descriptores en Ciencias de la Salud [Spanish translation of Medical Subject Headings (MeSH)], 2005. Sao Paulo: Latin American and Caribbean Center on Health Sciences Information. BIREME/PAHO/WHO, 2005.

Contact: Elenice de Castro; e-mail: elenice@brm.bireme.br

MTHMST2001 - Metathesaurus Version of Minimal Standard Terminology Digestive Endoscopy. Bethesda, MD: National Library of Medicine, 2001.

Contact: Jan Willis, National Library of Medicine, UMLS Support, 38A-4th fl, 8600 Rockville Pike, Bethesda, MD 20894; phone: 301 496-7715; e-mail: jwillis@nlm.nih.gov

---

MTHMSTFRE_2001 - Metathesaurus Version of Minimal Standard Terminology Digestive Endoscopy, French Translation. Bethesda, MD: National Library of Medicine, 2001.

***NOTE:** Now a CATEGORY 0.

Contact: Jan Willis, National Library of Medicine, UMLS Support, 38A-4th fl, 8600 Rockville Pike, Bethesda MD 20894; phone: 301 496-7715; e-mail: jwillis@nlm.nih.gov. NOTE: Users must also obtain rights to use the parent source.

MTHMSTITA_2001 - Metathesaurus Version of Minimal Standard Terminology Digestive Endoscopy, Italian Translation. Bethesda, MD: National Library of Medicine, 2001.

***NOTE:** Now a CATEGORY 0.

Contact: Jan Willis, National Library of Medicine, UMLS Support, 38A-4th fl, 8600 Rockville Pike, Bethesda MD 20894; phone: 301-496-7715; e-mail: jwillis@nlm.nih.gov

---

MTHPDQ2005 (updated) - Methathesaurus Forms of Physician Data Query (PDQ), 2005. Bethesda (MD): National Cancer Institute.

These terms were created by NLM to provide contextual information for PDQ terms.

Contact: Jan Willis, National Library of Medicine, UMLS Support, 38A-4th floor, 8600 Rockville Pike, Bethesda MD 20894; phone: 301-496-7715; email: jwillis@nlm.nih.gov

---

MTHSCT_2005_07_31 (updated) - Metathesaurus forms of SNOMED Clinical Terms. Bethesda (MD): National Library of Medicine, July 31, 2005.

to U.S. UMLS USERS

to Non-U.S. UMLS USERS

Contact: Jan Willis, National Library of Medicine, 38A-4th floor, 8600 Rockville Pike, Bethesda MD 20894; phone: 301-496-

7715; e-mail: jwillis@nlm.nih.gov

MTHSCTSPA_2005_04_30 (updated) - Metathesaurus forms of Spanish SNOMED Clinical Terms. Bethesda (MD): National Library of Medicine, April 30, 2005.

CATEGORY 4 RESTRICTIONS APPLY to U.S. UMLS USERS

CATEGORY 3 RESTRICTIONS APPLY to Non-U.S. UMLS USERS

These terms were created by NLM to provide contextual information for Spanish SNOMED Clinical Terms.

Contact: Jan Willis, National Library of Medicine, 38A-4th floor, 8600 Rockville Pike, Bethesda MD 20894; phone: 301-496-7715; e-mail: jwillis@nlm.nih.gov

---

NAN2004 - Sparks Ralph, Sheila, Craft-Rosenberg, Martha, Herdman, T. Heather, Lavin, Mary Ann, editors. NANDA nursing diagnoses: definitions and classification 2003-2004. Philadelphia: NANDA International, 2003.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Mary A. Lavin [lavinma@slu.edu]

NCBI2004_09_30 - NCBI Taxonomy. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Library of Medicine, National Center for Biotechnology Information, September 30, 2004. http://www.ncbi.nlm.nih.gov/Taxonomy/

Contact: http://www.ncbi.nlm.nih.gov/Taxonomy

NCI2004_11_17 (updated) - NCI Thesaurus. Bethesda (MD): National Cancer Institute, National Institutes of Health, November 2004. Contact: Sherri de Coronado, decorons@exchange.nih.gov, http://nci.nih.gov

Subset only.

Contact: Francis W. Hartel, PhD, Center for Bioinformatics, National Cancer Institute, 6116 Executive Blvd. Room 4019, Rockville MD 20892-8335; phone: 301-435-3869; email: hartel@mail.nih.gov

---

NCI-CTCAEV3 (new) - NCI modified Common Terminology Criteria for Adverse Events v3.0, Enterprise Vocabulary Service Project, National Cancer Institute. Bethesda, MD: National Cancer Institute, 2005.

Contact: M. Haber, National Cancer Institute, Building 6116, 6116 Exec Blvd, Room 3124, Rockville, MD; Phone: 301-594-9185; Fax: 301-480-8105; Email: mhaber@mail.nih.gov

---

NCISEER_1999 - NCI Surveillance, Epidemiology, and End Results (SEER) conversions between ICD-9-CM and ICD-10 neoplasm codes. National Cancer Institute, Bethesda, MD. Release Date: June 1999. URL: http://www-seer.ims.nci.nih.gov/Admin/ConvProgs/ Phone: 301-496-8510.

Contact: National Cancer Institute Bethesda, MD; phone: 301-496-8510

NDDF_2005_05_18 (updated) - National Drug Data File Plus Source Vocabulary 2005. San Bruno, CA: First DataBank, May 18, 2005.

CATEGORY 3 RESTRICTIONS APPLY

Contact: First DataBank Customer Support, 1111 Bayhill Drive, San Bruno, CA 94066; phone: 800-633-3453; e-mail: cs@firstdatabank.com

---

NDFRT_2004_01 - National Drug File - Reference Terminology, 2004_01. Washington, DC: U.S. Department of Veterans Affairs, Veterans Health Administration, January 2004.

Contact: Steven Brown; CPEP Office; 1310 24th Avenue S; Nashville, TN 37215; e-mail: Steven.Brown@msd.va.gov

---

NEU99 - Bowden, Douglas M., Martin, Richard F., Dubach, Joev G. Neuronames Brain Hierarchy. Seattle (WA): University of Washington, Primate Information Center, 1999. http://rprcsgi.rprc.washington.edu/neuronames/

CATEGORY 3 RESTRICTIONS APPLY

Contact: Douglas M. Bowden, M.D., Regional Primate Research Center, University of Washington, Box 357330, Seattle, WA 98195; e-mail: dmbowden@u.washington.edu

---

NIC99 - McCloskey, Joanne C., Bulechek, Gloria M., editors. NIC (Nursing Interventions Classification): Iowa Intervention Project. 2nd ed. St. Louis (MO): Mosby-Year Book, 1999.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Katherine Taylor Mosby-Year Book phone: (800)345-8738, ext. 7419 email: katherine.taylor@mosby.com

---

NLM-MED - National Library of Medicine (NLM) Medline Data. Bethesda (MD): National Library of Medicine.

Contact: Jan Willis, National Library of Medicine, 38A-4th floor, 8600 Rockville Pike, Bethesda MD 20894; phone: 301-496-7715; e-mail: jwillis@nlm.nih.gov

---

NOC97 - Johnson, Marion, Maas, Meridean, editors. Nursing Outcomes Classification (NOC): Iowa Outcomes Project. St.

Louis (MO): Mosby-Year Book, 1997.

Contact: Katherine Taylor, Mosby-Year Book; phone: (800)345-8738, ext. 7419; email: katherine.taylor@mosby.com

---

OMS94 - Martin, Karen S., Scheet, Nancy J. The Omaha System: Applications for Community Health Nursing. Philadelphia (PA): W.B. Saunders, 1992 (with 1994 corrections).

Contact: Karen S. Martin, RN, MSN, FAAN 2115 South 130th Street Omaha, Nebraska 68144; fax: (402)333-2091

PCDS97 - Ozbolt, Judy Grace. Patient Data Care Set (PCDS), Version 4.0, 1998.Contact: judy.ozbolt@mcmail.vanderbilt.edu; Vanderbilt University School of Nursing; 400-C Godchaux Hall; Nashville, TN 37240-0008; Telephone 615-343-3291

Contact: Judy Ozbolt, Vanderbilt School of Nursing 400-C Godchaux Hall Nashville TN 37240-0008; phone: (615)343-3291; email: judy.ozbolt@mcmail.vanderbilt.edu

---

PDQ2005 (updated) - PDQ. Bethesda (MD): National Cancer Institute, 2005.

Contact: Margaret Haber, NCI Building 6116 - 6116 Exec Blvd, Room 3124, Rockville, MD; telephone: 301-594-9185; Fax: 301-480-8105; Email: mhaber@mail.nih.gov

---

PNDS2002 (new) - Perioperative nursing data set : the perioperative nursing vocabulary, 2nd Ed. edited by Suzanne C. Beyea. / AORN; Denver, Colo. : AORN, 2002

Contact: Susan Kleinbeck, info@aorn.org

PPAC98 - Pharmacy Practice Activity Classification (PPAC). Version 1. Washington (DC): American Pharmaceutical Association, 1998.

Contact: Scott Antall, American Pharmaceutical Association - Academy of Pharmaceutical Research and Science; email: ssa@mail.aphanet.org; Academy of Pharmaceutical Research and Science, 2215 Constitution Avenue NW, Washington DC 20037-2985

PSY2004 (updated) - American Psychological Association. Thesaurus of psychological index terms/Lisa A. Gallagher, ed., 10th ed., 30th anniversary, 1974-2004. Washington, DC: American Psychological Association, c2005.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Lisa A. Gallagher, American Psychological Association, 750 First Stree, NE, Washingtion, DC 20002-4242; phone: 202-336-5726; email" LGallagher@apa.org

---

QMR96 - Quick Medical Reference (QMR). San Bruno (CA): First DataBank, 1997.

Contact: Quick Medical Reference, First Databank, 1111 Bayhill Drive San Bruno, CA 94066

RAM99 - QMR clinically related terms from Randolf A. Miller, 1999.

Contact: Dr. Randolph A. Miller (email: randolph.a.miller@vanderbilt.edu), Chair, Dept. of Biomedical Informatics, Vanderbilt University, 436 Eskind Biomedical Library, 2209 Garland Ave., Nashville TN 37232-8340

RCD99 - Clinical Terms Version 3 (CTV3) (Read Codes) (Q199): National Health Service National Coding and Classification Centre; March, 1999.

CATEGORY 3 RESTRICTIONS APPLY

Contact: NHS Information Authority Loughborough, Woodgate, Loughborough, Leicestershire LE11 2TG; fax: +44 (0) 1509 211611; e-mail: helpdesk3@nhsccc.exec.nhs.uk

---

RCDAE_1999 - American English equivalent of the Clinical Terms Version 3 (Q1, 1999), produced by NLM. Bethesda (MD): National Library of Medicine, UMLS project, 1999.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Jan Willis, National Library of Medicine, 38A-4th floor, 8600 Rockville Pike, Bethesda MD 20894; phone: 301-496-7715; e-mail: jwillis@nlm.nih.gov

---

RCDSA_1999 - American English equivalent of synthesized terms from the Clinical Terms Version 3 (Q1, 1999), produced by NLM. Bethesda (MD): National Library of Medicine, UMLS project, 1999.

CATEGORY 3 RESTRICTIONS APPLY

Contact: NHS Information Authority Loughborough, Woodgate, Loughborough, Leicestershire LE11 2TG; fax: +44 (0) 1509 211611; e-mail: helpdesk3@nhsccc.exec.nhs.uk

---

RCDSY_1999 - Synthesized Read terms (without initial bracketed letters) of the Clinical Terms Version 3 (Q1, 1999), produced by NLM. Bethesda (MD): National Library of Medicine, UMLS project, 1999.

CATEGORY 3 RESTRICTIONS APPLY

Contact: NHS Information Authority Loughborough, Woodgate, Loughborough, Leicestershire LE11 2TG; fax: +44 (0) 1509 211611; e-mail: helpdesk3@nhsccc.exec.nhs.uk

---

RXNORM_2005AC (updated) - RxNorm work done by NLM. National Library of Medicine (NLM). Bethesda (MD): National Library of Medicine, META2005AC Full Update 2005_09_14

This release contains concepts created by the National Library of Medicine which express the meaning of a drug name in a normalized form. These concepts relate the names of orderable medications to a dose form and the components of those medications. For further discussion, see the article at:

http://umlsinfo.nlm.nih.gov/RxNorm.html

Contact: Stuart Nelson, M.D., Head, MeSH Section; e-mail: nelson@nlm.nih.gov

---

SCTSPA_2005_04_30 (updated) - College of American Pathologists, SNOMED Clinical Terms, Spanish Language Edition, April 30, 2005. SNOMED International, 325 Waukegan Road, Northfield, IL 60093-2750. Phone: 800-323-4040 ext. 7700. Email: snomed@cap.org. URL: http://www.snomed.org

CATEGORY 4 RESTRICTIONS APPLY to U.S. UMLS USERS

CATEGORY 3 RESTRICTIONS APPLY to Non-U.S. UMLS USERS

Users are required to display the SNOMED copyright and trademark notice below if SNOMED information is used in their product(s).
"This publication includes SNOMED CT, a copyrighted work of the College of American Pathologists. ©2000, 2002 College of American Pathologists. This work is also protected by patent, U.S. Patent No. 6,438,533. SNOMED CT is used by permission of, and under license from, the College. SNOMED CT has been created by combining SNOMED RT and a computer based nomenclature and classification known as Clinical Terms Version 3, formerly known as Read Codes, Version 3, which was created on behalf of the U.K. Department of Health and is a crown copyright. SNOMED is a registered trademark of the College of American Pathologists."

Contact: SNOMED International, 325 Waukegan Road, Northfield, IL 60093-2750. Phone: 800-323-4040 ext. 7700. Email: snomed@cap.org. URL: http://www.snomed.org

---

SNM2 - Cote, Roger A., editor. Systematized Nomenclature of Medicine. 2nd ed. Skokie (IL): College of American Pathologists, 1979. SNOMED update, 1982. Skokie (IL): College of American Pathologists, 1982.

CATEGORY 4 RESTRICTIONS APPLY to U.S. UMLS USERS

CATEGORY 3 RESTRICTIONS APPLY to Non-U.S. UMLS USERS

Users are required to display the SNOMED copyright and trademark notice below if SNOMED information is used in their

product(s).

"This publication includes SNOMEDII, a copyrighted work of the College of American Pathologists. ©1979, 1980 College of American Pathologists. SNOMEDII is used by permission of, and under license from, the College. SNOMED is a registered trademark of the College of American Pathologists."

Contact: Jill Rosenthal/Naomi Siebert, College of American Pathologists, 325 Waukegan Road, Northfield, IL 60093-2750; fax: (847)832-8335; e-mail: jnrs@cap.org

---

SNMI98 - Cote, Roger A., editor. Systematized Nomenclature of Human and Veterinary Medicine: SNOMED International. Northfield (IL): College of American Pathologists; Schaumburg (IL): American Veterinary Medical Association, Version 3.5, 1998.

CATEGORY 4 RESTRICTIONS APPLY to U.S. UMLS USERS

CATEGORY 3 RESTRICTIONS APPLY to Non-U.S. UMLS USERS

Users are required to display the SNOMED copyright and trademark notice below if SNOMED information is used in their product(s).

"This publication includes SNOMED V3.5, a copyrighted work of the College of American Pathologists. ©1998 College of American Pathologists. SNOMED V3.5 is used by permission of, and under license from, the College. SNOMED is a registered trademark of the College of American Pathologists."

Contact: Jill Rosenthal/Naomi Siebert, College of American Pathologists, 325 Waukegan Road, Northfield, IL 60093-2750; fax: (847)832-8335; e-mail: jnrs@cap.org

---

SNOMEDCT_2005_07_31 (updated) - College of American Pathologists, SNOMED Clinical Terms. SNOMED International, 325 Waukegan Road, Northfield, IL 60093-2750. Phone: 800-323-4040 ext. 7700. Email: snomed@cap.org. Release Date: July 31, 2005. URL: http://www.snomed.org

CATEGORY 4 RESTRICTIONS APPLY to U.S. UMLS USERS

CATEGORY 3 RESTRICTIONS APPLY to Non-U.S. UMLS USERS

Users are required to display the SNOMED copyright and trademark notice below if SNOMED information is used in their product(s).

"This publication includes SNOMED CT, a copyrighted work of the College of American Pathologists. ©2000, 2002 College of American Pathologists. This work is also protected by patent, U.S. Patent No. 6,438,533. SNOMED CT is used by permission of, and under license from, the College. SNOMED CT has been created by combining SNOMED RT and a computer based nomenclature and classification known as Clinical Terms Version 3, formerly known as Read Codes, Version 3, which was created on behalf of the U.K. Department of Health and is a crown copyright. SNOMED is a registered trademark of the College of American Pathologists."

Contact: SNOMED International, 325 Waukegan Road, Northfield, IL 60093-2750. Phone: 800-323-4040 ext. 7700. Email: snomed@cap.org. URL: http://www.snomed.org

SPN2003 - Standard Product Nomenclature (SPN). Rockville, (MD); U.S. Food and Drug Administration, 2003

Contact: Stolber, Carey [Cstolber@ECRI.org]

SRC - UMLS Metathesaurus Source Terminologies. Bethesda, MD: National Library of Medicine.

Contact: Jan Willis, National Library of Medicine, UMLS Support, 38A-4th fl, 8600 Rockville Pike, Bethesda MD 20894; phone: 301 496-7715; e-mail: jwillis@nlm.nih.gov

---

ULT93 - Bell, Douglas. Ultrasound Structured Attribute Reporting (UltraSTAR). Boston (MA): Brigham & Womens Hospital, 1993.

CATEGORY 3 RESTRICTIONS APPLY

Contact: Robert Greenes, M.D., Ph.D., Brigham & Womens Hospital; Department of Radiology, 75 Francis Street, Boston MA 02115 e-mail: greenes@harvard.edu phone: (617) 732-6281

UMD2005 - Universal Medical Device Nomenclature System: Product Category Thesaurus. Plymouth Meeting (PA): ECRI, 2005.

CATEGORY 1 RESTRICTIONS APPLY

Contact: Elizabeth Richardson (erichard@ecri.org), Director of Database and Nomenclature Systems, ECRI, 5200 Butler Pike, Plymouth Meeting, PA 19462-1298; Phone: (610) 825-6000

USPMG_2004 (new) - United States Pharmacopeia (USP). Medicare Prescription Drug Benefit Model Guidelines: Drug Categories and Classes in Part D, 2004. http://www.usp.org/pdf/EN/mmg/comprehensiveDrugListing2004-12-31.pdf

Contact: http://www.usp.org/healthcareInfo/mmg/finalGuidelines.html

---

UWDA173 - University of Washington Digital Anatomist, (UWDA). Seattle (WA): University of Washinton, Version 1.7.3, March, 2003. Jose Mejino, M.D.; email: onard@biostr.washington.edu

Contact: Jose Mejino, M.D.; e-mail: onard@biostr.washington.edu; University of Washington Digital Anatomist Symbolic Knowledge Base, University of Washington Digital Anatomist Information System, Structural Informatics Group, Department of Biological Structure, University of Washington, Seattle WA 98195

VANDF_2005_03_23 (updated) - U.S. Department of Veterns Affairs, Veterans Health Administration National Drug File. Department of Veterans Affairs, Washington, DC. URL: http://www.vapbm.org/PBM/natform.htm

***NOTE:** Now a CATEGORY 0.

Contact: U.S. Department of Veterans Affairs, Veterans Health Administration, Washington DC; Steven Brown, CPEP office 1310 24th Avenue South, Nashville TN 37215; email: Steven.Brown@med.va.gov

---

WHO97 - WHO Adverse Drug Reaction Terminology (WHOART). Uppsala (Sweden): WHO Collaborating Centre for International Drug Monitoring, 1997.

CATEGORY 2 RESTRICTIONS APPLY

The Metathesaurus includes translations of WHO97 in:

- French (WHOFRE_1997),
- German (WHOGER_1997),
- Portuguese (WHOPOR_1997), and
- Spanish (WHOSPA_1997).

Contact: WHO Collaborating Centre for International Drug Monitoring, Stora Target 3, S-753 20 Uppsala, Sweden; fax: +46-18-656080

WHOFRE_1997 - WHO Adverse Drug Reaction Terminology (WHOART). French Translation. Uppsala (Sweden): WHO Collaborating Centre for International Drug Monitoring, 1997.

CATEGORY 2 RESTRICTIONS APPLY

Contact: WHO Collaborating Centre for International Drug Monitoring, Stora Target 3, S-753 20 Uppsala, Sweden; fax: +46-18-656080

WHOGER_1997 - WHO Adverse Drug Reaction Terminology (WHOART). German Translation. Uppsala (Sweden): WHO Collaborating Centre for International Drug Monitoring, 1997.

CATEGORY 2 RESTRICTIONS APPLY

Contact: WHO Collaborating Centre for International Drug Monitoring, Stora Target 3, S-753 20 Uppsala, Sweden; fax: +46-18-656080

WHOPOR_1997 - WHO Adverse Drug Reaction Terminology (WHOART). Portuguese Translation. Uppsala (Sweden): WHO Collaborating Centre for International Drug Monitoring, 1997.

CATEGORY 2 RESTRICTIONS APPLY

Contact: WHO Collaborating Centre for International Drug Monitoring, Stora Target 3, S-753 20 Uppsala, Sweden; fax: +46-18-656080

WHOSPA_1997 - WHO Adverse Drug Reaction Terminology (WHOART). Spanish Translation. Uppsala (Sweden): WHO Collaborating Centre for International Drug Monitoring, 1997.

CATEGORY 2 RESTRICTIONS APPLY

Contact: WHO Collaborating Centre for International Drug Monitoring, Stora Target 3, S-753 20 Uppsala, Sweden; fax: +46-18-656080

Table of Contents

APPENDIX B

METATHESAURUS® METADATA

B.0 Introduction

Appendix B provides details on Metathesaurus data referenced in Section 2 of this documentation.

B.1 Columns and Data Elements

Lists column and data element abbreviations, names and descriptions in Metathesaurus files in alphabetical order by abbreviation; includes field length and SQL92 datatype information. Provided for both Original Release and Rich Release Formats.

B.1.1 Columns and Data Elements in Rich Release Format (RRF)
B.1.2 Columns and Data Elements in Original Release Format (ORF)

B.2 Attribute Names

Lists attribute names and definitions in alphabetical order by abbreviation.

B.3 Abbreviations Used in Data Elements

Lists abbreviations and definitions of abbreviations used in Metathesaurus data elements, alphabetically by attribute type; includes relationship attributes.

B.4 Source Vocabularies

Lists alphabetically by source abbreviation vocabularies and classifications that are the sources of the concepts, terms and relationships in the Metathesaurus. HIPAA standard and CHI recommended vocabularies and code sets are noted.

B.5 Source and Term Types: Default Order of Precedence and Suppressibility

Lists sources and term types in default order of rank or precedence, used to determine referred names in the Metathesaurus, and notes the default suppressibility status assigned to each Source|Term Type.

B.6 Release Metadata

Provides an overview of the current Metatheasaurus, including summary counts of concepts, names, sources and languages.

B.1 Columns and Data Elements

All data elements in the Metathesaurus are described in this section. The data elements have been divided into Column Descriptions and Attribute Descriptions. The descriptions are arranged alphabetically by data element abbreviation.

Columns are described for Rich Release Format (RRF) in B.1.1, and for Original Release Format (ORF) in B.1.2.

B1.1 Columns and Data Elements in Rich Release Format (RRF)

Abbreviation	Description	File	Length of Value in characters	Average Length of Value in characters	SQL92 Datatype
ATN	Attribute name	MRSAT.RRF	2 - 24	8.85	varchar(50)
ATNL	Attribute name list for a source	MRSAB.RRF	0 - 826	28.90	varchar(1000)
ATUI	Unique identifier for attribute	MRDEF.RRF	10	10.00	varchar(10)
		MRSAT.RRF	10	10.00	varchar(10)
		MRSTY.RRF	10	10.00	varchar(10)
ATV	Attribute value	MRSAT.RRF	0 - 7903	6.81	varchar(4000)
AUI	Unique identifier for atom	MRCONSO. RRF	8	8.00	char(8)
			8	8.00	char(8)
		MRCXT.RRF	8	8.00	char(8)
		MRDEF.RRF	8	8.00	char(8)
		MRHIER.RRF			
AUI1	Unique identifier for first atom	MRREL.RRF	0 - 8	7.73	char(8)
		MRAUI.RRF	8	8.00	char(8)
		MRCOC.RRF	8	8.00	char(8)
AUI2	Unique identifier for second atom	MRREL.RRF	0 - 8	7.73	char(8)
		MRCOC.RRF	0 - 8	7.98	char(8)
		MRAUI.RRF	8	8.00	char(8)
		MRCXT.RRF	8	8.00	char(8)
AV	Average Length, Characters	MRCOLS.RRF	4 - 6	4.11	numeric(5,2)
BTS	Size in Bytes	MRFILES.RRF	1 - 11	7.16	integer
CENC	Character encoding of a source as specified by IANA	MRSAB.RRF	5	5.00	varchar(20)
CFR	CUI frequency for a source	MRSAB.RRF	0 - 6	3.79	integer
CHANGEKEY	CONCEPTSTATUS (if history relates to a SNOMED CT concept) or DESCRIPTIONSTATUS (if history relates to a SNOMED CT atom or "description")	MRHIST.RRF	13 - 17	15.77	varchar(1000)
CHANGETYPE	Source asserted code for type of change	MRHIST.RRF	1	1.00	varchar(1000)
CHANGEVAL	SNOMED CT CONCEPTSTATUS or DESCRIPTIONSTATUS value after the change took place	MRHIST.RRF	1 - 2	1.00	varchar(1000)
CLS	Number of columns	MRFILES.RRF	1 - 2	1.16	integer
COA	Attributes of co-occurrence	MRCOC.RRF	0 - 258	9.52	varchar(300)
CODE	Unique Identifier or code for string in source	MRSAT.RRF	0 - 30	6.01	varchar(50)
		MRCONSO.	1 - 30	7.57	varchar(50)
		RRF	1 - 30	7.88	varchar(50)
		MRCXT.RRF			
COF	Frequency of co-occurrence	MRCOC.RRF	0 - 5	1.10	integer
COL	Column or data element name	MRCOLS.RRF	2 - 11	3.82	varchar(20)
COT	Type of co-occurrence	MRCOC.RRF	1 - 3	1.10	varchar(3)

CUI	Unique identifier for concept	AMBIGLUI. RRF	8	8.00	char(8)
		AMBIGSUI. RRF	8	8.00	char(8)
		CHANGE/ MERGEDCUI. RRF	8	8.00	char(8)
		MRCONSO. RRF	8	8.00	char(8)
		MRCXT.RRF	8	8.00	char(8)
		MRDEF.RRF	8	8.00	char(8)
		MRHIER.RRF	8	8.00	char(8)
		MRHIST.RRF	8	8.00	char(8)
		MRSAT.RRF	8	8.00	char(8)
		MRSTY.RRF	8	8.00	char(8)
		MRXNS_ENG. RRF	8	8.00	char(8)
		MRXNW_ENG. RRF	8	8.00	char(8)
		MRXW_BAQ. RRF	8	8.00	char(8)
		MRXW_CZE. RRF	8	8.00	char(8)
		MRXW_DAN. RRF	8	8.00	char(8)
		MRXW_DUT. RRF	8	8.00	char(8)
		MRXW_ENG. RRF	8	8.00	char(8)
		MRXW_FIN. RRF			
		MRXW_FRE. RRF			
		MRXW_GER. RRF			
		MRXW_HEB. RRF			
		MRXW_HUN. RRF			
		MRXW_ITA. RRF			
		MRXW_JPN. RRF			
		MRXW_NOR. RRF			
		MRXW_POR. RRF			
		MRXW_RUS. RRF			
		MRXW_SPA. RRF			
		MRXW_SWE. RRF			
CUI1	Unique identifier for first concept	MRAUI.RRF	8	8.00	char(8)
		MRCOC.RRF	8	8.00	char(8)
		MRCUI.RRF	8	8.00	char(8)
		MRREL.RRF	8	8.00	char(8)

CUI2	Unique identifier for second concept	MRCUI.RRF	0 - 8	7.13	char(8)
		MRCOC.RRF	0 - 8	7.98	char(8)
		MRAUI.RRF	8	8.00	char(8)
		MRCXT.RRF	8	8.00	char(8)
		MRREL.RRF	8	8.00	char(8)
CURVER	Current Version flag	MRSAB.RRF	1	1.00	char(1)
CVF	Content view flag	MRCOC.RRF	0	0.00	integer
		MRCXT.RRF	0	0.00	integer
		MRDEF.RRF	0	0.00	integer
		MRHIER.RRF	0	0.00	integer
		MRHIST.RRF	0	0.00	integer
		MRMAP.RRF	0	0.00	integer
		MRREL.RRF	0	0.00	integer
		MRSAT.RRF	0	0.00	integer
		MRSMAP.RRF	0	0.00	integer
		MRSTY.RRF	0	0.00	integer
		MRCONSO. RRF	0 - 3	0.95	integer
CXL	Context member label, i.e., ANC for ancestor of this atom, CCP for the atom itself, SIB for sibling of this atom, CHD for child of this atom	MRCXT.RRF	3	3.00	char(3)
CXN	The context number if the atom has multiple contexts	MRCXT.RRF	1 - 4	1.90	integer
		MRHIER.RRF	1 - 4	1.91	integer
CXS	String for context member	MRCXT.RRF	1 - 1741	25.17	varchar(3000)
CXTY	Context type for a source	MRSAB.RRF	0 - 31	4.84	varchar(50)
DEF	Definition	MRDEF.RRF	1 - 5796	222.28	varchar(6000)
DES	Descriptive Name	MRCOLS.RRF	5 - 140	28.95	varchar(200)
		MRFILES.RRF	8 - 42	18.34	varchar(200)
DIR	Source asserted directionality flag	MRREL.RRF	0 - 1	0.47	varchar(1)
DOCKEY	Key to be documented	MRDOC.RRF	2 - 23	3.52	varchar(50)
DTY	SQL-92 data type for this column	MRCOLS.RRF	7 - 13	9.15	varchar(20)
EXPL	Detailed explanation	MRDOC.RRF	0 - 416	32.24	varchar(1000)
FIL	Physical FILENAME	MRCOLS.RRF	9 - 21	10.67	varchar(50)
		MRFILES.RRF	9 - 21	12.00	varchar(50)
FMT	Comma separated list of COL	MRFILES.RRF	7 - 190	33.32	varchar(300)
FROMEXPR	The expression that a mapping is mapped from.	MRMAP.RRF	2 - 9	7.81	varchar(4000)
		MRSMAP.RRF	2 - 9	7.81	varchar(4000)
FROMID	Metathesaurus identifier for mapped from	MRMAP.RRF	4 - 9	7.44	varchar(50)
FROMRES	Mapped from restriction.	MRMAP.RRF	0	0.00	varchar(4000)
FROMRULE	Rule for applying mapped from.	MRMAP.RRF	0	0.00	varchar(4000)
FROMSID	Source asserted identifier for mapped from	MRMAP.RRF	0 - 9	5.77	varchar(50)
FROMTYPE	The type of expression that a mapping is mapped from.	MRMAP.RRF	3 - 4	3.94	varchar(50)
		MRSMAP.RRF	3 - 4	3.94	varchar(50)
HCD	Source asserted hierarchical number or code of context member (if it exists)	MRCXT.RRF	0 - 43	0.10	varchar(50)
		MRHIER.RRF	0 - 43	0.11	varchar(50)
IMETA	Version of the Metathesaurus that a source was added	MRSAB.RRF	6	6.00	varchar(10)
ISPREF	Indicates whether AUI is preferred	MRCONSO. RRF	1	1.00	char(1)

LAT	Language of Term(s)	MRSAB.RRF	0 - 3	2.90	char(3)
		CHANGE/	3	3.00	char(3)
		DELETEDSUI.	3	3.00	char(3)
		RRF	3	3.00	char(3)
		MRCONSO.	3	3.00	char(3)
		RRF	3	3.00	char(3)
		MRXNS_ENG.	3	3.00	char(3)
		RRF	3	3.00	char(3)
		MRXNW_ENG.	3	3.00	char(3)
		RRF	3	3.00	char(3)
		MRXW_BAQ.	3	3.00	char(3)
		RRF	3	3.00	char(3)
		MRXW_CZE.	3	3.00	char(3)
		RRF	3	3.00	char(3)
		MRXW_DAN.	3	3.00	char(3)
		RRF	3	3.00	char(3)
		MRXW_DUT.	3	3.00	char(3)
		RRF	3	3.00	char(3)
		MRXW_ENG.	3	3.00	char(3)
		RRF	3	3.00	char(3)
		MRXW_FIN.	3	3.00	char(3)
		RRF	3	3.00	char(3)
		MRXW_FRE.			
		RRF			
		MRXW_GER.			
		RRF			
		MRXW_HEB.			
		RRF			
		MRXW_HUN.			
		RRF			
		MRXW_ITA.			
		RRF			
		MRXW_JPN.			
		RRF			
		MRXW_NOR.			
		RRF			
		MRXW_POR.			
		RRF			
		MRXW_RUS.			
		RRF			
		MRXW_SPA.			
		RRF			
		MRXW_SWE.			
		RRF			
LUI	Unique identifier for term	CHANGE/	0	0.00	char(8)
		MERGEDLUI.	0 - 8	6.07	char(8)
		RRF	8	8.00	char(8)
		MRSAT.RRF	8	8.00	char(8)
		AMBIGLUI.	8	8.00	char(8)
		RRF	8	8.00	char(8)
		MRCONSO.	8	8.00	char(8)
		RRF	8	8.00	char(8)
		MRXNS_ENG.	8	8.00	char(8)
		RRF	8	8.00	char(8)
		MRXNW_ENG.	8	8.00	char(8)
		RRF	8	8.00	char(8)
		MRXW_BAQ.	8	8.00	char(8)
		RRF	8	8.00	char(8)
		MRXW_CZE.	8	8.00	char(8)
		RRF	8	8.00	char(8)
		MRXW_DAN.	8	8.00	char(8)

		RRF	8	8.00	char(8)
		MRXW_DUT. RRF	8	8.00	char(8)
		MRXW_ENG. RRF	8	8.00	char(8)
		MRXW_FIN. RRF	8	8.00	char(8)
		MRXW_FRE. RRF	8	8.00	char(8)
		MRXW_GER. RRF			
		MRXW_HEB. RRF			
		MRXW_HUN. RRF			
		MRXW_ITA. RRF			
		MRXW_JPN. RRF			
		MRXW_NOR. RRF			
		MRXW_POR. RRF			
		MRXW_RUS. RRF			
		MRXW_SPA. RRF			
		MRXW_SWE. RRF			
MAPATN	Mapping attribute name (or future use).	MRMAP.RRF	0	0.00	varchar(20)
MAPATV	Mapping attribute value (for future use).	MRMAP.RRF	0	0.00	varchar(4000)
MAPID	Metathesaurus asserted identifier for mapping	MRMAP.RRF	10	10.00	varchar(50)
		MRSMAP.RRF	10	10.00	varchar(50)
MAPIN	Mapping in current subset	MRCUI.RRF	0 - 1	0.89	char(1)
		MRAUI.RRF	1	1.00	char(1)
MAPRANK	Ordering of mapping entries within a subset id	MRMAP.RRF	0 - 1	0.66	integer
MAPREASON	Reason for mapping	MRCUI.RRF	0 - 4	0.00	varchar(4000)
		MRAUI.RRF	4	4.00	varchar(4000)
MAPRES	Human readable restriction on when to apply mapping	MRMAP.RRF	0	0.00	varchar(4000)
MAPRULE	Machine processable rule for when to apply mapping	MRMAP.RRF	0	0.00	varchar(4000)
MAPSETCUI	CUI of the map set	MRMAP.RRF	8	8.00	char(8)
		MRSMAP.RRF	8	8.00	char(8)
MAPSETSAB	SAB of the map set	MRMAP.RRF	3 - 13	7.00	varchar(20)
		MRSMAP.RRF	3 - 13	7.00	varchar(20)
MAPSID	Source asserted identifier for mapping	MRMAP.RRF	0	0.00	varchar(50)
		MRSMAP.RRF	0	0.00	varchar(50)
MAPSUBSETID	Map sub set identifier	MRMAP.RRF	0 - 1	0.66	varchar(10)
MAPTYPE	Type of mapping	MRMAP.RRF	0 - 3	0.85	varchar(50)
MAX	Maximum Length	MRCOLS.RRF	1 - 4	1.37	integer
METAUI	Metathesaurus asserted unique identifier	MRSAT.RRF	0 - 9	7.01	varchar(50)
MIN	Minimum Length	MRCOLS.RRF	1 - 2	1.03	integer
NSTR	Normalized string	MRXNS_ENG. RRF	1 - 1630	33.21	varchar(3000)
NWD	Normalized word	MRXNW_ENG. RRF	1 - 80	6.38	varchar(100)

PAUI	Unique identifier for parent atom	MRHIER.RRF	0 - 8	8.00	char(8)
PCUI	Concept unique identifier in the previous Metathesaurus	CHANGE/ DELETEDCUI. RRF CHANGE/ MERGEDCUI. RRF	8 8	8.00 8.00	char(8) char(8)
PLUI	Lexical unique identifier in the previous Metathesaurus	CHANGE/ MERGEDLUI. RRF CHANGE/ DELETEDLUI. RRF	0 8	0.00 8.00	char(8) char(8)
PSTR	Preferred name in the previous Metathesaurus	CHANGE/ DELETEDSUI. RRF CHANGE/ DELETEDLUI. RRF CHANGE/ DELETEDCUI. RRF	1 - 940 1 - 940 2 - 735	28.93 30.72 72.10	varchar(3000) varchar(3000) varchar(3000)
PSUI	String unique identifier in the previous Metathesaurus	CHANGE/ DELETEDSUI. RRF	8	8.00	char(8)
PTR	Path to root	MRHIER.RRF	0 - 359	93.80	varchar(1000)
RANK	Termgroup ranking	MRCXT.RRF MRRANK.RRF	0 - 2 4	0.91 4.00	integer integer
RCUI	Unique identifier for root SRC concept	MRSAB.RRF	8	8.00	char(8)
REASON	Explanation of change, if present	MRHIST.RRF	0 - 48	2.87	varchar(1000)
REF	Documentation Section Number	MRCOLS.RRF	0	0.00	varchar(20)
REL	Relationship	MRAUI.RRF MRMAP.RRF MRSMAP.RRF MRCUI.RRF MRREL.RRF	0 0 - 2 0 - 2 2 - 3 2 - 3	0.00 1.88 1.88 2.11 2.54	varchar(4) varchar(4) varchar(4) varchar(4) varchar(4)
RELA	Relationship attribute	MRAUI.RRF MRCUI.RRF MRCXT.RRF MRHIER.RRF MRREL.RRF MRMAP.RRF MRSMAP.RRF	0 0 0 - 17 0 - 17 0 - 32 0 - 13 0 - 13	0.00 0.00 2.42 2.67 5.37 7.21 7.21	varchar(100) varchar(100) varchar(100) varchar(100) varchar(100) varchar(100) varchar(100)
RG	Relationship group	MRREL.RRF	0 - 3	0.12	varchar(10)
RMETA	Version of the Metathesaurus where a version is removed	MRSAB.RRF	0 - 6	0.12	varchar(10)
RSAB	Root source abbreviation	MRSAB.RRF	2 - 15	5.49	varchar(20)
RUI	Unique identifier for relationship	MRREL.RRF	9	9.00	varchar(10)
RWS	Number of rows	MRFILES.RRF	1 - 9	5.48	integer

SAB	Source abbreviation	MRDEF.RRF	2 - 10	3.09	varchar(20)
		MRRANK.RRF	2 - 15	4.87	varchar(20)
		MRREL.RRF	2 - 15	4.92	varchar(20)
		MRSAT.RRF	2 - 13	5.71	varchar(20)
		MRCONSO. RRF	2 - 15	5.72	varchar(20)
		MRCXT.RRF	2 - 9	6.73	varchar(20)
		MRHIER.RRF	2 - 9	6.98	varchar(20)
		MRCOC.RRF	3 - 5	3.01	varchar(20)
		MRHIST.RRF	8	8.00	varchar(20)
SABIN	Source in current subset	MRSAB.RRF	1	1.00	char(1)
SATUI	Source asserted attribute identifier	MRDEF.RRF	0	0.00	varchar(50)
		MRSAT.RRF	0	0.00	varchar(50)
SAUI	Source asserted atom identifier	MRCONSO. RRF	0 - 11	3.49	varchar(50)
SCC	Content contact info for a source	MRSAB.RRF	0 - 259	85.54	varchar(1000)
SCIT	Source citation	MRSAB.RRF	27 - 445	164.89	varchar(4000)
SCUI	Source asserted concept identifier	MRCONSO. RRF	0 - 24	4.32	varchar(50)
SDUI	Source asserted descriptor identifier	MRCONSO. RRF	0 - 13	1.93	varchar(50)
SF	Source Family	MRSAB.RRF	2 - 13	4.34	varchar(20)
SL	Source of relationship labels	MRREL.RRF	2 - 15	4.92	varchar(20)
SLC	License contact info for a source	MRSAB.RRF	0 - 295	137.99	varchar(1000)
SON	Source Official Name	MRSAB.RRF	10 - 145	48.57	varchar(3000)
SOURCEUI	Source asserted unique identifier	MRHIST.RRF	6 - 10	8.93	varchar(50)
SRL	Source Restriction Level	MRCONSO. RRF	1	1.00	integer
		MRSAB.RRF	1	1.00	integer
SRUI	Source attributed relationship identifier	MRREL.RRF	0 - 10	1.11	varchar(50)
SSN	Source short name	MRSAB.RRF	3 - 59	20.96	varchar(3000)
STN	Semantic type tree number	MRSTY.RRF	1 - 14	8.25	varchar(100)
STR	String	MRCONSO. RRF	1 - 1741	34.22	varchar(3000)
STT	String type	MRCONSO. RRF	2 - 3	2.01	varchar(3)
STY	Semantic type	MRSTY.RRF	4 - 41	18.52	varchar(50)
STYPE	The name of the column in MRCONSO.RRF or MRREL.RRF that contains the identifier to which the attribute is attached	MRSAT.RRF	3 - 4	3.61	varchar(50)
STYPE1	The name of the column in MRCONSO.RRF that contains the first identifier to which the relationship is attached	MRREL.RRF	3 - 4	3.31	varchar(50)
STYPE2	The name of the column in MRCONSO.RRF that contains the second identifier to which the relationship is attached	MRREL.RRF	3 - 4	3.31	varchar(50)

Col	Description	Source	Min-Max	Average	Type
SUI	Unique identifier for string	MRSAT.RRF	0 - 8	6.07	char(8)
		AMBIGSUI.RRF	8	8.00	char(8)
		MRCONSO.RRF	8	8.00	char(8)
		MRCXT.RRF	8	8.00	char(8)
		MRXNS_ENG.RRF	8	8.00	char(8)
		MRXNW_ENG.RRF	8	8.00	char(8)
		MRXW_BAQ.RRF	8	8.00	char(8)
		MRXW_CZE.RRF	8	8.00	char(8)
		MRXW_DAN.RRF	8	8.00	char(8)
		MRXW_DUT.RRF	8	8.00	char(8)
		MRXW_ENG.RRF	8	8.00	char(8)
		MRXW_FIN.RRF	8	8.00	char(8)
		MRXW_FRE.RRF	8	8.00	char(8)
		MRXW_GER.RRF			
		MRXW_HEB.RRF			
		MRXW_HUN.RRF			
		MRXW_ITA.RRF			
		MRXW_JPN.RRF			
		MRXW_NOR.RRF			
		MRXW_POR.RRF			
		MRXW_RUS.RRF			
		MRXW_SPA.RRF			
		MRXW_SWE.RRF			
SUPPRESS	Suppressible flag	MRCONSO.RRF	1	1.00	char(1)
		MRDEF.RRF	1	1.00	char(1)
		MRRANK.RRF	1	1.00	char(1)
		MRREL.RRF	1	1.00	char(1)
		MRSAT.RRF	1	1.00	char(1)
SVER	Release date or version number of a source	MRSAB.RRF	0 - 15	4.66	varchar(20)
		MRHIST.RRF	8	8.00	varchar(20)
TFR	Term frequency for a source	MRSAB.RRF	0 - 6	3.89	integer
TOEXPR	The expression that a mapping is mapped to.	MRMAP.RRF	0 - 242	6.82	varchar(4000)
		MRSMAP.RRF	0 - 242	6.82	varchar(4000)
TOID	Metathesaurus identifier for mapped to	MRMAP.RRF	0 - 10	5.50	varchar(50)
TORES	Mapped to restriction.	MRMAP.RRF	0	0.00	varchar(4000)
TORULE	Rule for applying mapped to.	MRMAP.RRF	0	0.00	varchar(4000)
TOSID	Source asserted identifier for mapped to	MRMAP.RRF	0 - 10	4.08	varchar(50)

TOTYPE	The type of expression that a mapping is mapped to.	MRMAP.RRF	0 - 18	4.41	varchar(50)
		MRSMAP.RRF	0 - 18	4.41	varchar(50)
TS	Term status	MRCONSO. RRF	1	1.00	char(1)
TTY	Term type in source	MRCONSO. RRF	2 - 4	2.14	varchar(20)
		MRRANK.RRF	2 - 4	2.15	varchar(20)
TTYL	Term type list for a source	MRSAB.RRF	0 - 88	11.22	varchar(100)
TUI	Unique identifier of Semantic type	MRSTY.RRF	4	4.00	char(4)
TYPE	Type of information	MRDOC.RRF	9 - 21	12.89	varchar(50)
VALUE	Value	MRDOC.RRF	0 - 55	8.25	varchar(1000)
VCUI	Unique identifier for versioned SRC concept	MRSAB.RRF	0 - 8	7.68	char(8)
VEND	Valid end date for a source	MRSAB.RRF	0	0.00	char(10)
VER	Last release version in which CUI1 was valid	MRAUI.RRF	6	6.00	varchar(10)
		MRCUI.RRF	6	6.00	varchar(10)
VSAB	Versioned source abbreviation	MRSAB.RRF	3 - 20	10.06	varchar(20)
VSTART	Valid start date for a source	MRSAB.RRF	0 - 10	0.14	char(10)
WD	Word in lower-case	MRXW_FIN. RRF	1 - 54	10.66	varchar(100)
			1 - 34	5.99	varchar(100)
		MRXW_JPN. RRF	1 - 19	6.12	varchar(100)
			1 - 80	6.17	varchar(100)
		MRXW_HEB. RRF	1 - 25	6.38	varchar(100)
			1 - 26	6.42	varchar(100)
		MRXW_ENG. RRF	1 - 46	6.63	varchar(100)
			1 - 38	6.78	varchar(100)
		MRXW_DAN. RRF	1 - 18	7.17	varchar(100)
			1 - 20	7.18	varchar(100)
		MRXW_NOR. RRF	1 - 38	7.34	varchar(100)
			1 - 33	7.60	varchar(100)
		MRXW_SPA. RRF	1 - 46	8.22	varchar(100)
			1 - 46	8.29	varchar(100)
		MRXW_POR. RRF	1 - 36	8.64	varchar(100)
			1 - 45	9.22	varchar(100)
		MRXW_BAQ. RRF	1 - 38	9.97	varchar(100)
		MRXW_HUN. RRF			
		MRXW_FRE. RRF			
		MRXW_ITA. RRF			
		MRXW_CZE. RRF			
		MRXW_DUT. RRF			
		MRXW_RUS. RRF			
		MRXW_GER. RRF			
		MRXW_SWE. RRF			
XC	Has Child	MRCXT.RRF	0 - 1	0.06	varchar(1)

B.1.2 Columns and Data Elements in Original Release Format (ORF)

Abbreviation	Description	File	Length of Value in characters	Average Length of Value in characters	SQL92 Datatype
ATN	Attribute name	MRSAT	2 - 24	8.15	varchar(20)
ATNL	Attribute name list for a source	MRSAB	0 - 826	28.90	varchar(1000)
ATV	Attribute value	MRSAT	0 - 7903	7.46	varchar(8000)
ATX	Associated expression	MRATX	5 - 242	35.84	varchar(300)
AV	Average Length, Characters	MRCOLS	4 - 6	4.10	numeric(5,2)
BTS	Size in Bytes	MRFILES	1 - 11	7.02	integer
CENC	Character encoding of a source as specified by IANA	MRSAB	5	5.00	varchar(20)
CFR	CUI frequency for a source	MRSAB	0 - 6	3.79	integer
CLS	Number of columns	MRFILES	1 - 2	1.05	integer
COA	Attributes of co-occurrence	MRCOC	0 - 258	9.52	varchar(300)
CODE	Unique Identifier or code for string in source	MRSAT MRSO MRCXT	0 - 30 1 - 30 1 - 30	6.64 7.57 7.88	varchar(50) varchar(50) varchar(50)
COF	Frequency of co-occurrence	MRCOC	1 - 5	1.10	integer
COL	Column or data element name	MRCOLS	2 - 6	3.13	varchar(10)
COT	Type of co-occurrence	MRCOC	1 - 3	1.10	varchar(3)
CREL	Relationship to retired CUI (CUI1) to current CUI (CUI2)	MRCUI	2 - 3	2.11	varchar(4)
CUI	Unique identifier for concept	AMBIG. LUI AMBIG. SUI CHANGE/ MERGED. CUI MRATX MRCON MRCXT MRDEF MRSAT MRSO MRSTY MRXNS. ENG MRXNW. ENG MRXW. BAQ MRXW. CZE MRXW. DAN MRXW. DUT MRXW. ENG MRXW. FIN MRXW. FRE MRXW. GER MRXW.	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00 8.00	char(8) char(8)

		HEB			
		MRXW. HUN			
		MRXW. ITA			
		MRXW. JPN			
		MRXW. NOR			
		MRXW. POR			
		MRXW. RUS			
		MRXW. SPA			
		MRXW. SWE			
CUI1	Unique identifier of first concept	MRCOC	8	8.00	char(8)
		MRCUI	8	8.00	char(8)
		MRREL	8	8.00	char(8)
CUI2	Unique identifier of second concept	MRCUI	0 - 8	7.13	char(8)
		MRCOC	0 - 8	7.98	char(8)
		MRCXT	8	8.00	char(8)
		MRREL	8	8.00	char(8)
CURVER	Current Version flag	MRSAB	1	1.00	char(1)
CXL	Context member label	MRCXT	3	3.00	char(3)
CXN	Context number	MRCXT	1 - 4	1.90	integer
CXS	String for context member	MRCXT	1 - 1741	25.17	varchar(1500)
CXTY	Context type for a source (as per section 2.3.2)	MRSAB	0 - 31	4.84	varchar(50)
DEF	Definition	MRDEF	1 - 5796	222.10	varchar(6000)
DES	Descriptive Name	MRCOLS	4 - 56	26.14	varchar(100)
		MRFILES	7 - 42	17.92	varchar(100)
DTY	SQL-92 data type for this column	MRCOLS	7 - 13	8.66	varchar(20)
FIL	Physical FILENAME	MRCOLS	4 - 18	7.19	varchar(50)
		MRFILES	4 - 18	8.32	varchar(50)
FMT	Comma separated list of COL	MRFILES	7 - 112	20.98	varchar(150)
HCD	Hierarchical number or code of context member	MRCXT	0 - 43	0.10	varchar(50)
IMETA	Version of the Metathesaurus that a source was added	MRSAB	6	6.00	varchar(10)
LAT	Language of Term(s)	MRSAB	0 - 3	2.90	char(3)
		CHANGE/	3	3.00	char(3)
		DELETED.	3	3.00	char(3)
		SUI	3	3.00	char(3)
		MRCON	3	3.00	char(3)
		MRXNS. ENG	3	3.00	char(3)
		MRXNW. ENG	3	3.00	char(3)
		MRXW. BAQ	3	3.00	char(3)
		MRXW. CZE	3	3.00	char(3)
		MRXW. DAN	3	3.00	char(3)
		MRXW. DUT	3	3.00	char(3)
		MRXW. ENG	3	3.00	char(3)
		MRXW.	3	3.00	char(3)

		FIN	3	3.00	char(3)
		MRXW. FRE	3	3.00	char(3)
		MRXW. GER			
		MRXW. HEB			
		MRXW. HUN			
		MRXW. ITA			
		MRXW. JPN			
		MRXW. NOR			
		MRXW. POR			
		MRXW. RUS			
		MRXW. SPA			
		MRXW. SWE			
LRL	Least Restriction Level	MRCON	1	1.00	integer
LUI	Unique identifier for term	CHANGE/	0	0.00	char(8)
		MERGED.	0 - 8	6.70	char(8)
		LUI	8	8.00	char(8)
		MRSAT	8	8.00	char(8)
		AMBIG.	8	8.00	char(8)
		LUI	8	8.00	char(8)
		MRCON	8	8.00	char(8)
		MRSO	8	8.00	char(8)
		MRXNS.	8	8.00	char(8)
		ENG	8	8.00	char(8)
		MRXNW.	8	8.00	char(8)
		ENG	8	8.00	char(8)
		MRXW.	8	8.00	char(8)
		BAQ	8	8.00	char(8)
		MRXW.	8	8.00	char(8)
		CZE	8	8.00	char(8)
		MRXW.	8	8.00	char(8)
		DAN	8	8.00	char(8)
		MRXW.	8	8.00	char(8)
		DUT	8	8.00	char(8)
		MRXW.	8	8.00	char(8)
		ENG	8	8.00	char(8)
		MRXW.	8	8.00	char(8)
		FIN	8	8.00	char(8)
		MRXW. FRE			
		MRXW. GER			
		MRXW. HEB			
		MRXW. HUN			
		MRXW. ITA			
		MRXW. JPN			
		MRXW.			

144

		NOR MRXW. POR MRXW. RUS MRXW. SPA MRXW. SWE			
MAPIN	Mapping in current subset	MRCUI	0 - 1	0.89	char(1)
MAX	Maximum Length	MRCOLS	1 - 4	1.38	integer
MEND	Metathesaurus end date for a source	MRSAB	0	0.00	char(8)
MG	Machine generated and unverified indicator	MRREL	0	0.00	varchar(1)
MIN	Minimum Length	MRCOLS	1 - 2	1.00	integer
MSTART	Metathesaurus start date for a source	MRSAB	0 - 10	0.14	char(10)
NSTR	Normalized string	MRXNS. ENG	1 - 1630	33.21	varchar(1500)
NWD	Normalized word	MRXNW. ENG	1 - 80	6.38	varchar(100)
PCUI	Concept unique identifier in the previous Metathesaurus	CHANGE/ DELETED. CUI CHANGE/ MERGED. CUI	8 8	8.00 8.00	char(8) char(8)
PLUI	Lexical unique identifier in the previous Metathesaurus	CHANGE/ MERGED. LUI CHANGE/ DELETED. LUI	0 8	0.00 8.00	char(8) char(8)
PSTR	Preferred name in the previous Metathesaurus	CHANGE/ DELETED. SUI CHANGE/ DELETED. LUI CHANGE/ DELETED. CUI	1 - 940 1 - 940 2 - 735	28.93 30.72 72.10	varchar(3000) varchar(3000) varchar(3000)
PSUI	String unique identifier in the previous Metathesaurus	CHANGE/ DELETED. SUI	8	8.00	char(8)
RANK	Termgroup ranking	MRRANK	4	4.00	integer
RCUI	Unique identifier for root SRC concept	MRSAB	8	8.00	char(8)
REF	Documentation Section Number	MRCOLS	0	0.00	varchar(20)
REL	Relationship	MRATX MRREL	2 2 - 3	2.00 2.59	varchar(3) varchar(3)
RELA	Relationship attribute	MRCXT MRREL	0 - 17 0 - 32	2.42 4.68	varchar(100) varchar(100)
RMETA	Version of the Metathesaurus where a version is removed	MRSAB	0 - 6	0.12	varchar(10)
RNK	Rank	MRCXT	0 - 2	0.91	integer
RSAB	Root source abbreviation	MRSAB	2 - 15	5.49	varchar(20)
RWS	Number of rows	MRFILES	1 - 9	5.42	integer

SAB	Source abbreviation	MRDEF	2 - 10	3.09	varchar(20)
		MRREL	2 - 15	4.77	varchar(20)
		MRRANK	2 - 15	4.87	varchar(20)
		MRSAT	2 - 13	5.45	varchar(20)
		MRSO	2 - 15	5.72	varchar(20)
		MRCXT	2 - 9	6.73	varchar(20)
		MRATX	3	3.00	varchar(20)
SABIN	Source in current subset	MRSAB	1	1.00	char(1)
SCC	Content contact info for a source	MRSAB	0 - 259	85.54	varchar(1000)
SF	Source Family	MRSAB	2 - 13	4.34	varchar(20)
SL	Source of relationship labels	MRREL	2 - 15	4.77	varchar(20)
SLC	License contact info for a source	MRSAB	0 - 295	137.99	varchar(1000)
SOC	Source of co-occurrence information	MRCOC	3 - 5	3.01	varchar(20)
SON	Source Official Name	MRSAB	10 - 145	48.57	varchar(1500)
SRL	Source Restriction Level	MRSAB	1	1.00	integer
		MRSO	1	1.00	integer
STR	String	MRCON	1 - 1741	35.52	varchar(1500)
STT	String type	MRCON	2 - 3	2.01	varchar(3)
STY	Semantic type	MRSTY	4 - 41	18.52	varchar(50)
SUI	Unique identifier for string	MRSAT	0 - 8	6.70	char(8)
		AMBIG. SUI	8	8.00	char(8)
		MRCON	8	8.00	char(8)
		MRCXT	8	8.00	char(8)
		MRSO	8	8.00	char(8)
		MRXNS. ENG	8	8.00	char(8)
		MRXNW. ENG	8	8.00	char(8)
		MRXW. BAQ	8	8.00	char(8)
		MRXW. CZE	8	8.00	char(8)
		MRXW. DAN	8	8.00	char(8)
		MRXW. DUT	8	8.00	char(8)
		MRXW. ENG	8	8.00	char(8)
		MRXW. FIN	8	8.00	char(8)
		MRXW. FRE	8	8.00	char(8)
		MRXW. GER			
		MRXW. HEB			
		MRXW. HUN			
		MRXW. ITA			
		MRXW. JPN			
		MRXW. NOR			
		MRXW. POR			
		MRXW.			

Col Abbr	Description	Source	Min/Max Length	Average Length	SQL Data Type
		RUS MRXW. SPA MRXW. SWE			
SUPRES	Suppressible flag	MRRANK	1	1.00	char(1)
SVER	Release date or version number of a source	MRSAB	0 - 15	4.66	varchar(20)
TFR	Term frequency for a source	MRSAB	0 - 6	3.89	integer
TS	Term status	MRCON	1	1.00	char(1)
TTY	Term type in source	MRSO	2 - 4	2.14	varchar(20)
		MRRANK	2 - 4	2.15	varchar(20)
TTYL	Term type list for a source	MRSAB	0 - 88	11.22	varchar(100)
TUI	Unique identifier of Semantic type	MRSTY	4	4.00	char(4)
VCUI	Unique identifier for versioned SRC concept	MRSAB	0 - 8	7.68	char(8)
VER	Last release version in which CUI1 was valid	MRCUI	6	6.00	varchar(10)
VSAB	Versioned source abbreviation	MRSAB	3 - 20	10.06	varchar(20)
WD	Word in lower-case	MRXW. FIN	1 - 54	10.66	varchar(100)
		MRXW. JPN	1 - 34	5.99	varchar(100)
		MRXW. HEB	1 - 19	6.12	varchar(100)
		MRXW. ENG	1 - 80	6.17	varchar(100)
		MRXW. DAN	1 - 25	6.38	varchar(100)
		MRXW. NOR	1 - 26	6.42	varchar(100)
		MRXW. SPA	1 - 46	6.63	varchar(100)
		MRXW. POR	1 - 38	6.78	varchar(100)
		MRXW. BAQ	1 - 18	7.17	varchar(100)
		MRXW. HUN	1 - 20	7.18	varchar(100)
		MRXW. FRE	1 - 38	7.34	varchar(100)
		MRXW. ITA	1 - 33	7.60	varchar(100)
		MRXW. CZE	1 - 46	8.22	varchar(100)
		MRXW. DUT	1 - 46	8.29	varchar(100)
		MRXW. RUS	1 - 36	8.64	varchar(100)
		MRXW. GER	1 - 45	9.22	varchar(100)
		MRXW. SWE	1 - 38	9.97	varchar(100)
XC	Has Child	MRCXT	0 - 1	0.06	varchar(1)

B.2 Attribute Names

ATN (Attribute Name)	
AN	MeSH Annotation - an informative MeSH note written primarily for indexers or catalogers that may also be useful in explaining the use of a MeSH term to online searchers.
AQL	MeSH Allowable Qualifier - list of allowable qualifier abbreviations for MeSH main headings (e.g. AA, CL, CS, DF, DU, IM, I,P ME, PK)
AQ	SNOMED CT "allowable qualifier" attribute for representing certain relationships (those having a characteristic type of "Qualifier" and a refinability of "Mandatory") which indicate one of several allowable types of qualifiers, such as laterality or severity, that a concept may have
ASSOCABNORMAL_CELL	NCI role Disease_May_Have_Abnormal_Cell
ASSOCCELLORIGIN	Disease may have normal cell origin
ASSOCCYTOGENETICS	Disease may have cytogenetics
ASSOCFINDING	Disease may have finding
ASSOCIATEDDISEASE	Disease may have associated diseases
ASSOCMOLECTRAIT	Disease may have molecular abnormality
ASSOCTISSUEORIGIN	Disease may have normal tissue origin
ATC	Alternative Billing Codes Tree Code - actual alternative medicine billing alphabetic code string. Treenumber for context in MRCXT.
BIOCARTAID	BioCarta ID
BIOMARKERTYPE	Gene is biomarker type
CASRN	CAS registry
CCF	Canonical Clinical Problem Statement System (CCPSS) frequency - the number of times a CCPSS term appears in a patient record.
CCI	ICD-9-CM code(s) clusters in a Clinical Classifications Software (CCS) category - individual ICD-9-CM codes (or ranges of such codes) classified into CCS categories.
CFR	Code of Federal Regulation Number (e.g. 862.3220, 892.1610)
CHARACTERISTICTYPE	SNOMED CT indication of whether a relationship specifies a defining characteristic of the source concept or a possible qualification of that Concept.
COMPONENTHISTORY	SNOMED CT history of a single instance of a change to a source data element in a particular version of SNOMED CT.
CONCEPTSTATUS	SNOMED CT status which indicates whether a concept is in active use and, if not, indicates the reason it is inactive.
CONSIDER	Codes to consider before definitive diagnosis
CPA	CPT Short Description - CPT abbreviated procedure description (e.g. ANESTH)
CPF	CPT Full Procedure - complete text of the CPT full procedure, in cases where the CPT2002 term in the "STR" field of MRCON has been trimmed from its original form.
CRIT_INTERACT_WITH	Critical interaction with
CTV3ID	The Read Code for a SNOMED CT concept taken from the United Kingdom?s Clinical Terms Version 3 terminology.
CX	MeSH Consider Also Note - other word roots or prefixes that should be consulted for concepts related to this MeSH concept, e.g., the value for "Heart" is "consider also terms at cardi- and myocardi-".
DATE_CREATED	Date created
DATE_FIRST_PUBLISHED	Date first published
DATE_LAST_MODIFIED	Date last modified
DA	Metathesaurus Date of entry - YYYYMMDD, e.g., 19920830 - date of entry of the concept into the Metathesaurus.
DCSA	Controlled Substance Act designation code (e.g. 4)
DC	MeSH Descriptor class - type of MeSH term the concept name represents.
DDFA	Drug Doseform Abbreviation (e.g. SOLN)

DDF	Drug Doseform (e.g. chewable tablet)
DEFCHARACTERISTICS	Defining characteristics
DESCRIPTIONSTATUS	SNOMED CT description status which indicates whether a description (concept name) is in active use and, if not, the reason it is inactive.
DESCRIPTIONTYPE	SNOMED CT term type, indicating whether the term is the Preferred Term, Synonym or the Fully Specified Name for the associated concept.
DHJC	Multum HCPCS J-code Multum clinical drugs linked to HCPCS J-codes where applicable (e.g. J7507)
DID	Descriptor Identifier
DISPLAY_NAME	Display name
DIV	NCBI Division/Phyla (e.g. DIV[NCBI]Viruses)
DPC	Multum Pregnancy Hazard Classification Code assigned to Multum clinical drugs. (e.g. X, D)
DQ	MeSH Date Qualifier Established YYYYMMDD - date the qualifier became available for indexing MEDLARS citations.
DRTA	Drug Route of Administration Abbreviation (e.g. INJ)
DRT	Drug Route of Administration (e.g. Injection (systemic))
DST	Drug Strength (e.g. 0.01%, 0.02 MG, 0.02 MG/ML)
DS	MeSH Descriptor Sort Version - form needed for proper sequencing of the concept name, if the name could not be sequenced properly by the sort algorithms used in the MeSH publications (e.g. MC Antithrombin III DS Antithrombin 03)
DX	MeSH Date major descriptor established YYYYMMDD - first day of the Index Medicus publication month in which the descriptor (in any form) was available for searching as a major descriptor.
EC	MeSH Entry combination - an invalid MeSH main heading/subheading combination that is a cross reference to a single MeSH main heading or a main heading/subheading combination that should be used in its place.
EV	MeSH Entry term abbreviation - a short form for a MeSH entry term or cross reference used primarily in MEDLINE record creation and maintenance.
EXCLUDESABNORMALCELL	Disease excludes abnormal cell
EXCLUDESCELLORIGIN	Disease excludes normal cell origin
EXCLUDESCTYOABNORM	Disease excludes cytogenic abnormality
EXCLUDESFINDINGS	Disease excludes findings
EXCLUDESMOLECTRAIT	Disease excludes molecular abnormality
EXCLUDESPRIMARYSITE	Disease excludes primary anatomic site
EXCLUDE_DI_CHECK	Exclude drug interaction check
EXPRESSEDINTISSUE	Gene product expressed in tissue
EZ	Enzyme Commission Number - International Union of Biochemists Enzyme Commission number for an enzyme concept.
FOUNDINORGANISM	Gene found in organism
FROMRSAB	Root source abbreviation for the "from" identifiers of a map set
FROMVSAB	Versioned source abbreviation for the "from" identifiers of a map set
FR	MeSH Frequency
FX	MeSH MH Mapping - maps a MeSH MH to a 'See Related' MH.
GENBANKACCESSION	GenBank accession number
GENEENCODESPRODUCT	Gene encodes product
GOANNOTATION	GO annotation
GO_Comment	GO comment
GO_Namespace	GO namespace
GO_Subset	GO subset
GROUPURL	URL for corresponding health topic groups on MedlinePlus

GXR	GO Cross Reference to external databases (e.g. MetaCyc:TRNA-CHARGING-PWY)
HAB	HCPCS abbreviation (short form)
HAC	HCPCS action code - code denoting the change made to a procedure or modifier code within the HCPCS system.
HAD	HCPCS Action Effective Date - effective date of action to a procedure or modifier code.
HAQ	HCPCS Anesthesia Base Unit Quantity - base unit represents the level of intensity for anesthesia procedure services that reflects all activities except time.
HASABNORMALCELL	Disease has abnormal cells
HASABNORMALITY	Gene has abnormality
HASASSOCIATEDANATOMY	Disease has associated anatomy
HASASSOCIATEDDISEASE	Disease has associated disease
HASASSOCIATEDSITE	Disease has associated anatomic site
HASBIOCHEMFUNCTION	Disease has biochemical function
HASCELLORIGIN	Disease has normal cell origin
HASCHEMCLASS	Disease has chemical class
HASCONCEPTSTATUS	Has concept status
HASCYTOGENETICS	Disease has cytogenetic abnormality
HASDOMAINORMOTIF	Disease has domain or motif
HASFINDING	Disease has finding
HASMETASTATICSITE	Disease has metastatic site
HASMOLECTRAIT	Disease has molecular abnormality
HASORGANISMSOURCE	Disease has organism source
HASPRIMARYSITE	Disease has primary anatomic site
HASTISSUEORIGIN	Disease has normal tissue origin
HAT	HL7 "appliesTo" concept code property
HBT	HCPCS Berenson-Eggers Type of Service Code - BETOS for the procedure code based on generally agreed upon clinically meaningful groupings of procedures and services.
HCC	HCPCS Coverage Code - code denoting Medicare coverage status. There are two subelements separated by "=".
HCDI	HL7 Code Instance: Number to disambiguate case differences (Default 0)
HCD	HCPCS Code Added Date - year the HCPCS code was added to the HCFA Common Procedure Coding System.
HCI	HL7 Codesystem Id: Formal OID of the specific coding system
HCO	HL7 Comment
HCS	HL7 Concept Status: Status of concept code (P - proposed, A - active, R - retired)
HDS	HL7 Designation Sequence: Designation id within concept
HFN	HL7 Codesystem fullname: Official name of the coding system
HHA	HL7 "howApplies" concept code property
HID	HL7 internal Id: Internal identifier of codeSystemId/conceptCode. More than one conceptCode/codeInstance may map to the same internalId
HIR	HCPCS Coverage Issues Manual Reference Section Number - number identifying the Reference Section of the Coverage Issues Manual.
HIT	HL7 Interpretation: German Interpretation according to the German HL/ user group
HLC	HCPCS Lab Certification Code - code used to classify laboratory procedures according to the specialty certification categories listed by CMS(formerly HCFA).
HLIR	HL7 "inverseRelationship" concept code property
HMP	HCPCS Multiple Pricing Indicator Code - code used to identify instances where a procedure could be priced.
HMR	HCPCS Medicare Carriers Manual reference section number - number identifying a section of the Medicare Carriers Manual

HM	MeSH Heading Mapped To - heading mapped to attribute in C-MeSH containing repeating (MH or MH/SH) elements (e.g. HM = PYRROLIDINONES, HM = *TARTRATES, HM = ESTRONE/* analogs & derivatives)
HN	History Note - for MeSH history notes, the year when the current form of the MeSH term was established as a major and/or minor descriptor.
HOD	HL7 OID
HOI	HL7 "openIssue" concept code property
HPD	HCPCS ACD payment group effective date - date the procedure is assigned to the ASC payment group.
HPG	HCPCS ASC payment group code which represents the dollar amount of the facility charge payable by Medicare for the procedure.
HPI	HCPCS Pricing Indicator Code - used to identify the appropriate methodology for developing unique pricing amounts under Part B.
HPL	HL7 Preferred for Language: Yes means that this designation is the primary designation for the given language
HPN	HCPCS processing note number identifying the processing note contained in Appendix A of the HCPCS Manual.
HRI	HL7 Codesystem Release Id: Release identifier record of of last update of this code system
HSE	HL7 Section: Chapter
HSNU	HL7 Sort Number: Number for sorting the values according to the official document
HSN	HCPCS Statute Number identifying statute reference for coverage or noncoverage of procedure or service.
HST	HL7 Codesystem Type: One of "I" (internal), "E" (external) or "EI" (external but internally maintained)
HTA	HL7 Anchor: HTML-Reference to the Standard
HTD	HCPCS Termination Date - last date for which a procedure or code may be used by Medicare Providers.
HTG	HL7 Generate: Yes/No
HTS	HCPCS Type of Service Code - carrier assigned HCFA Type of Service which describes the particular kind(s) of service represented by the procedure code.
HTT	HL7 Table Type: Number specifying the type of this table (user defined, HL7 defined, ..)
HUAL	HUGO aliases
HUDA	HUGO date approved
HUDC	HUGO date changed
HUDM	HUGO date modified
HUS	HL7 Usage: How is this data element used (Original, Added, Deleted)
HUXR	HUGO cross reference
HXR	HCPCS Cross reference code - an explicit reference crosswalking a deleted code or a code that is not valid for Medicare to a valid current code (or range of codes).
IAA	ICD10AM Abbreviated Descriptor
IAC	ICD10AM Australian Code - Contains a flag "1" for codes that have been exclusively developed in Australia.
IAD	ICD10AM Effective from - indicates the date that the code is effective from.
IAH	ICD10AM Age Edit, higher limit - minimum age is expressed as a three digit field.
IAL	ICD10AM Age Edit, lower limit - minimum age is expressed as a three digit field.
IAN	ICD10AM Annotation Note - a + (Dagger) denotes a code describing the etiology or underlying cause of the disease; an * (asterisk) denotes a code describing the manifestation of a disease
IAR	ICD10AM Relationship Note - code embedded in the term, may have a * or +.
IAS	ICD10AM Sex Edit flag to indicate whether the code is valid for a particular sex. Sex flags are 1 (male) or 2 (female).

IAT	ICD10AM Sex Edit Type - all sex flagged codes are followed by a "sex edit type" flag. The sex edit type flags are 1 (fatal) or 2 (warning).
IAY	ICD10AM Age Edit Type - all age flagged codes are followed by an age edit type flag. The age edit type flags are 1 (fatal) or 2 (warning).
ICA	ICD Additional Codes Note - an ICD-9-CM instruction to signal the coder that an additional code should be used if the information is available to provide a more complete picture of that diagnoses or procedure.
ICC	ICD Codes Also Note - an ICD-9-CM instruction to signal the coder to code additional information.
ICE	ICD Entry Term (e.g. Diarrhea:{dysenteric; epidemic}; Infectious diarrheal disease)
ICF	ICD Fifth-digit Code Note - instruction to coders indicating use of a fifth-digit code.
ICN	ICD Note - instruction providing additional coding information.
ICS	ICD Short Form - 25-character version of the code.
IDNR	Original ID number for source file
II	MeSH Indexing Information - for MeSH chemical terms (Term Type=NM), MeSH headings that may be relevant to articles that are also assigned the NM term.
INCHROMOSOMALLOCATION	Gene in chromosomal location
INCLUDED_MESH_UIS	Used for MoA concepts where multiple MeSH descriptors were merged into one NDF-RT concept
INC	ICD10AM Neoplasm code (e.g. C34.-)
INDCODE	IND code
INITIALCAPITALSTATUS	SNOMED CT capital status, indicating whether the capitalization of the first character of the SNOMED CT term is significant.
INTERVENTION_DESCRIPTION	Intervention description
INTERVENTION_SEQUENCE	Intervention sequence number
INTERVENTION_TEXT	Intervention text
IPX	ICD10 code related to an ICPC code - a + indicates that the ICD10 code is broader than the ICPC code; a - indicates that the ICD10 code is narrower than the ICPC code.
ISELEMENTINPATHWAY	Gene product is element in pathway
ISGRADE	Disease is grade
ISPRIMITIVE	SNOMED CT indicator of whether concept is primitive or fully defined by its defining characteristics.
ISSTAGE	Disease is stage
JXR	OMIM code that is related to Jablonski term.
KEGGID	KEGG ID
LAC	ASTM Code - E1238-94 code for a test in a LOINC name.
LAL	LOINC Answerlist - list of answers for results that are reportable from a multiple choice list, e.g., the answers for the term DISPOSITION OF BLOOD PACK are GIVEN; PARTIALLY GIVEN;DISCARDED.
LANGUAGECODE	SNOMED CT string identifying a language and, if appropriate, a dialect in which this description is valid.
LAST_REVIEWED	Last reviewed
LCA	LOINC ACSSYM field - LOINC chemical name synonyms, alternative names and chemical formulae from the Chemical Abstract Society.
LCB	LOINC Chemical base name from the Chemical Abstract Society.
LCC	LOINC CDC Code - code from the Centers for Disease Control Complexity file that maps a laboratory test to the instruments used to perform the test. The code is at the analyte level, not at the test instrument level.
LCD	LOINC CDISC common tests
LCI	LOINC Molecular structure ID, usually a Chemical Abstract Society number.

LCL	LOINC Class - arbitrary classification of terms in LOINC designed to assist LOINC development and to group related observations together (e.g. ABXBACT = Antibiotic susceptibility)
LCN	LOINC Classtype - 1 = Laboratory class; 2 = Clinical class
LCR	LOINC Reason for Change - a brief explanation of the change made to a LOINC term.
LCS	LOINC Depreciated or superseded status - an indicator that a LOINC term is no longer to be used. The term that should now be used will appear in the LMP element.
LCT	LOINC Change Type Code - type of change made to a LOINC term.
LDE	LOINC DEEDS_CD - Data Elements for Emergency Department Systems Codes (CDC). This field contains the DEEDS code value which maps to the LOINC code in question.
LEA	LOINC Example Answers - for some tests and measurements, LOINC has supplied examples of valid answers. These values differ from those in the ANSWERLIST field because that details possible values for nominal scale terms.
LEC	LOINC Analyte Code - EUCLIDES code for the analyte which is the first subpart of the first part of a LOINC name.
LEVEL	Specifies the type of DRUG_KIND concepts. Values are "VA Class," "Ingredient," "Ingredient Strength," "Orderable Product".
LFO	LOINC Formula - regression equation details for many OB.US calculated terms.
LFR	French name for LOINC term supplied by Centre Suisse de Controle de Qualite. This field contains extended characters and will not transfer to 7-bit systems.
LGC	GPI Code, GPI Code Total - for drugs, this field contains a map to the Medispan GPI codes, a hierarchical system of classifying pharmaceutical products. For a few products, a simple one-to-one mapping with a GIP code was not possible. In these cases, all applicable GPI codes are contained in this field, separated by semicolons.
LGR	German Name for LOINC Term - supplied by Centre Suisse de Controle de Qualite; contains extended characters and will not transfer correctly to 7-bit systems
LIC	IUPAC Code code for the component, kind of property, and system in a LOINC name. Note that most IUPAC codes assume that the component is measured in substance concentration, e.g., moles. The IUPAC code for substance concentration is applied to mass concentration in LOINC, because IUPAC has no code for the mass concentration variant, which is more commonly used in the U.S.
LIR	Italian Name for LOINC Term - supplied by Centre Suisse de Controle de Qualite; contains extended characters and will not transfer correctly to 7-bit systems
LIU	IUPAC Analyte code - contains the chemical abstract service number or the enzyme nomenclature number for the chemical components for chemicals and/or enzymes. These were also contributed by IUPAC.
LLR	LOINC Date Last Changed YYYYMMDD - date the LOINC term was last changed.
LMC	LOINC Metpath Code representing the LOINC name at MetPath laboratories.
LMM	Molecular weights - contains the molecular weights of many chemical moieties when they are provided.
LMP	LOINC Map to Code of the term that has superseded a term with a LCS value of DEL.
LMT	LOINC MULTUM_CD - maps to Multum Inc. database of codes for drugs.
LNC	LOINC NAACCR_ID - LOINC terms mapped to North American Association of Central Cancer Registries Identification
LNE	LOINC CODE_TABLE
LOCUSID	Locus ID
LOINC_COMPONENT	LOINC component
LOINC_METHOD_TYP	LOINC method type
LOINC_PROPERTY	LOINC property
LOINC_SCALE_TYP	LOINC scale type
LOINC_SYSTEM	LOINC system

LOINC_TIME_ASPECT	LOINC time aspect
LOR	LOINC ORDER_OBS field. Defines term as order only, observation only, or both. Values are: BOTH OBSERVATION ORDER. A fourth category, Subset, is used for terms that are subsets of a panel but do not represent a package that is known to be orderable.
LPL	LOINC Panel Elements
LQS	Survey Question Source
LQT	Survey Question Text
LRF	Reference - contains references to medical literature, product announcements, or other written sources of information on the test or measurement described by the LOINC record.
LRN2	Related names 2
LRN	LOINC related name - Previously was released as a RN string from LOINC; converted from a term type in 2002AD (e.g. AMIKIN)
LSC	Code for a SNOMED International laboratory procedure name that is related to (usually broader than) the LOINC term.
LSN	LOINC short name
LSP	LOINC Species code
LSR	Root of a set of LOINC codes - currently used for claims attachments. Yes in this field signifies that this record is the root of a set of LOINC codes.
LSU	SUBMITED_UNITS
LT	Indicates if a chemicals or medical device is a tradename (present in older versions of the Metathesaurus and was discontinued, then brought back starting in 2002AD)
LUN	LOINC Typical Units - typical units in which the observation is recorded.
LUR	Units required when used as OBX segment - a Y/N field that indicates that units are required when this LOINC is included as an OBX segment in a HIPAA attachment
MAPFROMCOMPLEXITY	Indicates the complexity of "from" expressions used in a map set
MAPFROMEXHAUSTIVE	Indicates whether or not the "from" source of a map set is completely mapped
MAPSETCOMPLEXITY	Indicates the overall complexity of a map set
MAPSETGRAMMAR	Grammar used to describe boolean expressions used in a map set
MAPSETID	Identifier for a map set
MAPSETNAME	Official name of a map set
MAPSETREALMID	Identifier of a "Realm" to which a source is mapped, within which this cross mapping table is applicable. Used in cases where Realm specific business rules or guidelines alter the acceptable mappings. Realm is the same as used in SNOMED CT subsets. It includes a four character ISO6523 identifier followed by an optional series of concatenated subdivision codes defined by the registered organization.
MAPSETRSAB	Root source abbreviation for a map set
MAPSETRULETYPE	Indicates the types of rules used in a map set and cross map targets to which a source is mapped.
MAPSETSCHEMEID	Standard identifier for the scheme to which a map set belongs. This may be an International Coding Scheme Identifier (ISO7826) or an Object Identifier (OID) used as specified by HL7.
MAPSETSCHEMENAME	Full name of the target scheme in a map set.
MAPSETSCHEMEVERSION	Version number of the target scheme (as published by the issuing organization) in a map set.
MAPSETSEPARATORCODE	XML entity code (for example, "&#7c;" to represent the vertical-bar character) for the character used as a separator between the individual codes in the target codes field in a map set.
MAPSETTYPE	Indicates the nature of a map set. Its value is map set specific. It can be used to indicate the inclusion of one to one, one to many and choices of maps.
MAPSETVSAB	Versioned source abbreviation for a map set

MAPSETXRTARGETID	Map set target identifier used bfor XR mappings.
MAPTOCOMPLEXITY	Indicates the complexity of "to" expressions in a map set
MAPTOEXHAUSTIVE	Indicates whether or not the "to" source is completely mapped
MDA	MeSH date of entry YYYYMMDD - date the term was added to the MeSH file, which is prior to the date the term became available for indexing and searching MEDLARS citations. Terms that have been part of MeSH for many years may have no value in this element.
MEA	In NOC, the "measurement scale" used for a particular outcome.
MED1901	Medline citation counts from articles dated 1901.
MED1902	Medline citation counts from articles dated 1902.
MED1903	Medline citation counts from articles dated 1903.
MED1904	Medline citation counts from articles dated 1904.
MED1905	Medline citation counts from articles dated 1905.
MED1906	Medline citation counts from articles dated 1906.
MED1907	Medline citation counts from articles dated 1907.
MED1908	Medline citation counts from articles dated 1908.
MED1909	Medline citation counts from articles dated 1909.
MED1910	Medline citation counts from articles dated 1910.
MED1911	Medline citation counts from articles dated 1911.
MED1912	Medline citation counts from articles dated 1912.
MED1913	Medline citation counts from articles dated 1913.
MED1914	Medline citation counts from articles dated 1914.
MED1915	Medline citation counts from articles dated 1915.
MED1916	Medline citation counts from articles dated 1916.
MED1917	Medline citation counts from articles dated 1917.
MED1918	Medline citation counts from articles dated 1918.
MED1919	Medline citation counts from articles dated 1919.
MED1920	Medline citation counts from articles dated 1920.
MED1921	Medline citation counts from articles dated 1921.
MED1922	Medline citation counts from articles dated 1922.
MED1923	Medline citation counts from articles dated 1923.
MED1924	Medline citation counts from articles dated 1924.
MED1925	Medline citation counts from articles dated 1925.
MED1926	Medline citation counts from articles dated 1926.
MED1927	Medline citation counts from articles dated 1927.
MED1928	Medline citation counts from articles dated 1928.
MED1929	Medline citation counts from articles dated 1929.
MED1930	Medline citation counts from articles dated 1930.
MED1931	Medline citation counts from articles dated 1931.
MED1932	Medline citation counts from articles dated 1932.
MED1933	Medline citation counts from articles dated 1933.
MED1934	Medline citation counts from articles dated 1934.
MED1935	Medline citation counts from articles dated 1935.
MED1936	Medline citation counts from articles dated 1936.
MED1937	Medline citation counts from articles dated 1937.
MED1938	Medline citation counts from articles dated 1938.
MED1939	Medline citation counts from articles dated 1939.
MED1940	Medline citation counts from articles dated 1940.
MED1941	Medline citation counts from articles dated 1941.
MED1942	Medline citation counts from articles dated 1942.

MED1943	Medline citation counts from articles dated 1943.
MED1944	Medline citation counts from articles dated 1944.
MED1945	Medline citation counts from articles dated 1945.
MED1946	Medline citation counts from articles dated 1946.
MED1947	Medline citation counts from articles dated 1947.
MED1948	Medline citation counts from articles dated 1948.
MED1949	Medline citation counts from articles dated 1949.
MED1950	Medline citation counts from articles dated 1950.
MED1951	Medline citation counts from articles dated 1951.
MED1952	Medline citation counts from articles dated 1952.
MED1953	Medline citation counts from articles dated 1953.
MED1954	Medline citation counts from articles dated 1954.
MED1955	Medline citation counts from articles dated 1955.
MED1956	Medline citation counts from articles dated 1956.
MED1957	Medline citation counts from articles dated 1957.
MED1958	Medline citation counts from articles dated 1958.
MED1959	Medline citation counts from articles dated 1959.
MED1960	Medline citation counts from articles dated 1960.
MED1961	Medline citation counts from articles dated 1961.
MED1962	Medline citation counts from articles dated 1962.
MED1963	Medline citation counts from articles dated 1963.
MED1964	Medline citation counts from articles dated 1964.
MED1965	Medline citation counts from articles dated 1965.
MED1966	Medline citation counts from articles dated 1966.
MED1967	Medline citation counts from articles dated 1967.
MED1968	Medline citation counts from articles dated 1968.
MED1969	Medline citation counts from articles dated 1969.
MED1970	Medline citation counts from articles dated 1970.
MED1971	Medline citation counts from articles dated 1971.
MED1972	Medline citation counts from articles dated 1972.
MED1973	Medline citation counts from articles dated 1973.
MED1974	Medline citation counts from articles dated 1974.
MED1975	Medline citation counts from articles dated 1975.
MED1976	Medline citation counts from articles dated 1976.
MED1977	Medline citation counts from articles dated 1977.
MED1978	Medline citation counts from articles dated 1978.
MED1979	Medline citation counts from articles dated 1979.
MED1980	Medline citation counts from articles dated 1980.
MED1981	Medline citation counts from articles dated 1981.
MED1982	Medline citation counts from articles dated 1982.
MED1983	Medline citation counts from articles dated 1983.
MED1984	Medline citation counts from articles dated 1984.
MED1985	Medline citation counts from articles dated 1985.
MED1986	Medline citation counts from articles dated 1986.
MED1987	Medline citation counts from articles dated 1987.
MED1988	Medline citation counts from articles dated 1988.
MED1989	Medline citation counts from articles dated 1989.
MED1990	Medline citation counts from articles dated 1990.
MED1991	Medline citation counts from articles dated 1991.
MED1992	Medline citation counts from articles dated 1992.

MED1993	Medline citation counts from articles dated 1993.
MED1994	Medline citation counts from articles dated 1994.
MED1995	Medline citation counts from articles dated 1995.
MED1996	Medline citation counts from articles dated 1996.
MED1997	Medline citation counts from articles dated 1997.
MED1998	Medline citation counts from articles dated 1998.
MED1999	Medline citation counts from articles dated 1999.
MED2000	Medline citation counts from articles dated 2000.
MED2001	Medline citation counts from articles dated 2001.
MED2002	Medline citation counts from articles dated 2002.
MED2003	Medline citation counts from articles dated 2003.
MED2004	Medline citation counts from articles dated 2004.
MED2005	Medline citation counts from articles dated 2005.
MENU_PARENT	Link information to Term Menu parent. This may not always be the same as the TermParent. Element will not exist for top level menus, e.g. cancer.
MENU_TYPE	Values are Clinical Trials--Cancer Type, Clinical Trials--Drug, or Cancer Information.
MESH_UI	MeSH UI
MISO	MedDRA Serial Code International SOC Sort Order Digit (01-26)
MMR	MeSH revision date YYYYMMDD - date of the last major revision to the term's MeSH record.
MN	MeSH hierarchical number for the concept in the MeSH tree structures. This number also appears in the HCD subelement of the REL and CXT elements.
MPS	MedDRA primary SOC (PTs may have multiple treepositions, but each has a primary soc)
MR	Major revision date YYYYMMDD - date the Metathesaurus entry for the concept underwent any revision in content.
MSA	MedDra abbreviation (either SOC or Special Search Category)
MSC	Minimal Standard (Terminology) Class
MSP	SPN Medical Specialty Panel (responsible for reviewing the product).
MUI	MeSH Unique Concept Identifier (MUI) assigned by NLM. Each concept in MeSH is given a MUI. (e.g. M0001333)
MXR	MedDRA cross reference to WHOART, COSTART, or ICD9-CM (e.g.: MXR
NAF	NIC atoms - complete text of the NIC full intervention, in cases where the NIC99 term in the "STR" field of MRCON has been trimmed from its original form (due to length or to extraneous information at the end of the text).
NAT	Neuronames Anatomy Type - indication of the type of anatomy represented by a Neuronames concept name.
NA	Neuronames Abbreviation - short abbreviation for a concept name in the Neuronames thesaurus.
NC_NAME	NCI concept name
NDC	National Drug Code corresponding to a clinical drug (e.g. 66109-ABD-00)
NDF_GCN_SEQ_NUMBER	NDF Generic Code Number (GCN) Sequence Number
NDF_STRENGTH	NDF strength
NDF_TRANSMIT_TO_CMOP	NDF Transmit to Consolidated Mail Outpatient Pharmacy (CMOP)
NDF_UNITS	NDF units
NFI	National formulary indicator - "YES" or "NO" indicating whether a drug is in the VA's National Formulary
NF_INACTIVATE	National Formulary Inactivation Date - the date a drug was removed from the VA's National Formulary
NF_NAME	National Formulary Name

NH	NonHuman Flag - single character: Y. An indication that the concept does not apply to human beings, used only when the concept's Semantic type(s) could imply the contrary. For example, the concept BEAK and CLAW are assigned the Semantic type" Body Part, Organ, or Organ Component", but do not apply to human beings.
NSCCODE	NSC code
NSR	Neuronames Species Restriction - indication that a Neuronames concept applies only to humans or only to macaques. Most Neuronames concepts apply to both and have no species restriction.
NST	Normalized strength and units for drugs with one active ingredient (e.g. 769 MG)
OL	MeSH Online Note - information helpful to online searchers of MEDLINE, especially when the history of a term or cross-reference has implications for online searching. This is a potential source of useful information for rules for search interface programs.
OMIMNUMBER	OMIM number
ORIG_CODE	Original code associated with this string
ORIG_SOURCE	Original versioned source associated with this string
ORIG_TTY	Original term type
ORIG_VSAB	Original versioned source abbreviation
OUTCOME_EXAMPLES	Outcome examples
OUTCOME_INDICATORS	Outcome indicators
OUTCOME_STATEMENT	Outcome statement
PARTOF	Part of
PA	Pharmacologic Action of MeSH main headings (MH) for drugs and supplementary concept names (NM). The information in this element is also represented by an "isa" relationship between the MH or NM concept and the MeSH concept name for the class of drugs with a particular pharmacologic action.
PCL	Pharmacy Practice Activity Classification (PPAC) Category - all terms are assigned to one of five categories, which connote their hierarchy.
PDC	SPN Product Device Class (level of CDRH regulation: class 1, 2, or 3).
PID	Legacy PDQ ID
PI	MeSH heading or heading/subheading combination(s) followed by a date range in parentheses (YYYY).
PLAYSROLEINBIOPROCESS	Gene product plays role in biological process
PLAYSROLEINPROCESS	Gene product plays role in process
PLR	Pharmacy Practice Activity Classification (PPAC) Last Revision Date (Format: M/DD/ YY time)
PM	Public MeSH note - combines key information from the HN and PI elements in a format that is printed in the MeSH publications.
PRC	Product Third Party Review Code from SPN.
PREFERREDNAME	Preferred name
PRN	VA print/label name
PRODUCTHASABNORMALITY	Gene product has abnormality
PRODUCTISBIOMARKERTYPE	Gene product is biomarker type
PTR	SPN Product Tier (level of CDRH triage: 1, 2, 3, or E{xempt}).
PXC	PDQ Protocol Exclusion Criteria - terms with type "exclusion criteria," which may be indexed on protocol records to identify conditions that exclude a patient from eligibility.
PYR	PsychInfo year designation
QA	MeSH Topical Qualifier Abbreviation - for MeSH subheadings (Term Type=TQ), an abbreviation that may be used in place of the full text of the abbreviation in searching on NLM's system and possibly on other systems offering NLM data.
QE	MeSH Qualifier Entry Version - short form for a MeSH qualifier.

QS	MeSH Qualifier Sort Version - form of the subheading needed for proper alphabetic sequencing when the subheading cannot be sequenced properly by the sort algorithms used in the MeSH publications.
RANK	NCBI Rank (e.g. RANK[NCBI]species)
REFINABILITY	Indicates whether it is possible or mandatory to refine the atom when this relationship is used as a template for clinical data entry.
REF	List of bibliographic citations related to a given nursing intervention or nursing outcome.
RELATEDFACTORS	Related factors
REMARKS	Specific coding instructions
RID	Read Codes Term_id - identifier assigned to a Read term, used in referring to the term in the Read file structure, and may be used in clinical information systems.
RISKFACTORS	Risk factors
RN	Registry Number - series of numbers and hyphens (any leading zeros in an RN are dropped) or a series of numbers and periods, preceded by EC
RR	The Chemical Abstracts Registry numbers for salts, optical isomers, or isotope-labeled versions of the concept followed by the relationship of this RR to the RN (in parentheses.) Applies to chemicals only. These numbers can be used as links to information in a number of chemical and toxicological databases.
RXAUI	RxNorm atom identifier for the atom it is connected to
RXCUI	RxNorm concept identifier for the atom this is connected to
RXNORM_SCDC	RxNorm SCDC
RXNORM_SCD	RxNorm SCD
RXN_OBSOLETED	Date the RxNorm atom became obsolete
RXO	Drug Description from NDF - R for Rx and O for OTC
SB	SNOMED International subset indicator - valid values: * =can code using two T codes or G code for laterality B Bethesda system (Morphology); IC= ICDO (Oncology) related; N=Nursing; N* =Nursing, provisional; U=Ultrastructure (Morphology); V= Veterinary= V* =Veterinary AND can code using two T codes or G code for laterality (Topography)
SHF	SNOMED Hospital Formulary Code - the American Hospital Formulary Code for a chemical contained in SNOMED International (e.g. 84:24:12)
SIC	SNOMED ICD9CM Reference - the ICD9CM code or codes listed as relevant to the meaning of the concept in SNOMED International.
SID	Secondary GO ID (e.g. GO:0020034)
SIGNIF_INTERACT_WITH	Significant interaction with
SMX	SNOMED Multiaxial coding - an alphanumeric string that includes hyphens, parentheses, and sometimes ellipses.
SNGL_OR_MULT_SRC_PRD	Single or multi-source product
SNOMEDID	SNOMED RT identifier for a SNOMED CT concept
SOS	Scope Statement Two subelements: 1. Abbreviation of the source of the statement; 2. Variable length string with alpha character, punctuation, and in some cases diacritics
SRC	MeSH Literature source of chemical name - a citation to an article in a journal indexed for MEDLINE in which the chemical has been identified. (Note: Not to be confused with source abbreviation of SRC)
ST	Concept Attributes Status - valid values: R Reviewed, U Unreviewed
SUBSETCONTEXTID	SNOMED CT identifier of a subset
SUBSETLANGUAGECODE	SNOMED CT identifier of a language and, if appropriate, a dialect to which a subset applies.
SUBSETMEMBER	Indicates the subset(s) to which an atom belongs and its status (e.g., active, inactive) in the subset(s).
SUBSETORIGINALID	SNOMED CT identifier for the first version of the subset on which this subset is based.

SUBSETREALMID	"Realm" or context in which a SNOMED CT subset is used, e.g., nation, speciality, institution, etc. to which it applies. A four-character ISO6523 identifier followed by an optional series of concatenated subdivisions codes defined by the registered organization.
SUBSETTYPE	Single digit integer indicating the nature of a SNOMED CT subset and the type of SNOMED CT component that may be a member of the subset. The meaning of the non-zero values can be found in the SNOMED CT documentation; a value of 0 is not defined as a type.
SUBSETVERSION	SNOMED CT identifier for the version of a subset. An integer increased for each revised release.
SWP	Swiss Protein Number
TARGETSCHEMEID	Identifier for the target scheme in the map set. This may be an International Coding Scheme Identifier (ISO7826) or an Object Identifier (OID) used as specified by HL7.
TERMUI	Term unique identifier
TH	MeSH Thesaurus ID - identifies thesauri other than MeSH in which the MeSH heading or cross-reference is included.
TORSAB	Root source abbreviation for the "to" identifiers of a map set
TOVSAB	Versioned source abbreviation for the "to" identifiers of a map set
TRN	Trade Name of Product
TYPE	Multum Medical Supply Category (e.g. natural supplements)
UMLSAUI	UMLS Atom Unique Identifier
UMLSCUI	UMLS CUI
UMLSMAPSETSEPARATOR	The character used in the UMLS Metathesaurus as a separator between the individual codes in the target codes field of the cross map targets to which a source is mapped.
UMLSRELA	UMLS relationship attribute
UMLSREL	The UMLS Metathesaurus REL relationship (SY, CHD, NT, BT, RT) assigned to SNOMED CT relationship identifiers.
UNII_CODE	FDA Unique Ingredient Identifier (UNII) Code
URL	URL for corresponding health topic on MedlinePlus
UWT	A semantic type provided from terms from the University of Washington Digital Anatomist
VAC	VA Class - the code of an NDF/HT drug class name (e.g. COD LIVER OIL (CHERRY FLAVOR))
VA_CLASS_NAME	VA class name
VA_DISPENSE_UNIT	VA dispense unit
VA_GENERIC_NAME	VA generic name
VMO	VA CMOP (central mail - order pharmacy) ID

COA (Co-Occurrence Attribute)	
<>	No MeSH subheading (qualifier)
AA	analogs & derivatives
AB	abnormalities
AD	administration & dosage
AE	adverse effects
AG	agonists
AH	anatomy & histology
AI	antagonists & inhibitors
AN	analysis
AO	in adolescence
AU	in adulthood
BI	biosynthesis

BL	blood
BS	blood supply
CF	cerebrospinal fluid
CH	chemistry
CI	chemically induced
CL	classification
CN	congenital
CO	complications
CS	chemical synthesis
CT	contraindications
CY	cytology
DE	drug effects
DF	deficiency
DH	diet therapy
DI	diagnosis
DT	drug therapy
DU	diagnostic use
EC	economics
ED	education
EH	ethnology
EM	embryology
EN	enzymology
EP	epidemiology
ES	ethics
ET	etiology
GD	growth & development
GE	genetics
HI	history
IC	in infancy & childhood
IM	immunology
IN	injuries
IP	isolation & purification
IR	innervation
IS	instrumentation
LJ	legislation & jurisprudence
MA	manpower
ME	metabolism
MI	microbiology
MO	mortality
MT	methods
MY	in middle age
NU	nursing
OA	in old age
OG	organization & administration
PA	pathology
PC	prevention & control
PD	pharmacology
PH	physiology
PK	pharmacokinetics
PO	poisoning

PP	physiopathology
PR	in pregnancy
PS	parasitology
PX	psychology
PY	pathogenicity
RA	radiography
RE	radiation effects
RH	rehabilitation
RI	radionuclide imaging
RT	radiotherapy
SC	secondary
SD	supply & distribution
SE	secretion
SN	statistics & numerical data
ST	standards
SU	surgery
TD	trends
TH	therapy
TM	transmission
TO	toxicity
TR	transplantation
TU	therapeutic use
UL	ultrastructure
UR	urine
US	ultrasonography
UT	utilization
VE	veterinary
VI	virology

B.3 Abbreviations Used in Data Elements

COT (Type of Co-Occurrence)	
KN	negative association in Knowledge Base, e.g., a finding that is inconsistent with a disease.
KP	positive association in Knowledge Base
LQB	second concept is qualified by the first (a MeSH topical qualifier) in citations to the published literature
LQ	second concept occurs as a MeSH topical qualifier of the first in citations to the published literature. Where CUI2 is not present, the count of citations of CUI1 with no MeSH qualifiers is reported.
L	Co-occurrence of primary or main subject headings in citations to the published literature
MP	Co-occurrence of modifier and problem within a patient record
PP	Co-occurrence of two problems within a patient record

FROMTYPE (Type of Expression from Which a Mapping is Mapped)

AUI	Atom identifier
BOOLEAN_EXPRESSION	Boolean expression of strings or identifiers
CODE	Unique Identifier or code for string in source
CUI	Concept unique identifier
SAUI	Source asserted atom unique identifier
SCUI	Source asserted concept unique identifier
SDUI	Source asserted descriptor identifier

LAT (Language of Terms)	
BAQ	Basque
CZE	Czech
DAN	Danish
DUT	Dutch
ENG	English
FIN	Finnish
FRE	French
GER	German
HEB	Hebrew
HUN	Hungarian
ITA	Italian
JPN	Japanese
NOR	Norwegian
POR	Portuguese
RUS	Russian
SPA	Spanish
SWE	Swedish

MAPATN	
	Empty attribute name

REL (Relationship)	
AQ	Allowed qualifier
CHD	has child relationship in a Metathesaurus source vocabulary
DEL	Deleted concept
PAR	has parent relationship in a Metathesaurus source vocabulary
QB	can be qualified by.
RB	has a broader relationship
RL	the relationship is similar or "alike". the two concepts are similar or "alike". In the current edition of the Metathesaurus, most relationships with this attribute are mappings provided by a source, named in SAB and SL; hence concepts linked by this relationship may be synonymous, i.e. self-referential: CUI1 = CUI2. In previous releases, some MeSH Supplementary Concept relationships were represented in this way.
RN	has a narrower relationship
RO	has relationship other than synonymous, narrower, or broader
RQ	related and possibly synonymous.
RU	Related, unspecified
SIB	has sibling relationship in a Metathesaurus source vocabulary.
SUBX	Concept removed from current subset
SY	source asserted synonymy.

XR	Not related
	Empty relationship

RELA (Relationship Attribute)	
access_instrument_of	Access instrument of
access_of	Access of
active_ingredient_of	Active ingredient of
actual_outcome_of	Actual outcome of
adjectival_form_of	Adjectival form of
adjustment_of	Adjustment of
affected_by	Affected by
affects	Affects
analyzed_by	Analyzed by
analyzes	Analyzes
approach_of	Approach of
associated_disease	Associated disease
associated_finding_of	Associated finding of
associated_genetic_condition	Associated genetic condition
associated_morphology_of	Associated morphology of
associated_procedure_of	Associated procedure of
associated_with	Associated with
branch_of	Branch of
british_form_of	British form of
causative_agent_of	Causative agent of
cause_of	Cause of
challenge_of	Challenge of
classified_as	Classified as
classifies	Classifies
clinically_associated_with	Clinically associated with
clinically_similar	Clinically similar
co-occurs_with	Co-ocurrs with
component_of	Component of
conceptual_part_of	Conceptual part of
consists_of	Consists of
constitutes	Constitutes
contained_in	Contained in
contains	Contains
contraindicated_with	Contraindicated with
course_of	Course of
ddx	dxx
default_mapped_from	Default mapped from
default_mapped_to	Default mapped to
definitional_manifestation_of	Definitional manifestation of
degree_of	Degree of
diagnosed_by	Diagnosed by
diagnoses	Diagnoses
direct_device_of	Direct device of
direct_morphology_of	Direct morphology of
direct_procedure_site_of	Direct procedure site of

direct_substance_of	Direct substance of
divisor_of	Divisor of
dose_form_of	Dose form of
drug_contraindicated_for	Drug contraindicated for
due_to	Due to
encoded_by_gene	Encoded by gene
encodes_gene_product	Encodes gene product
episodicity_of	Episodicity of
evaluation_of	Evaluation of
exhibited_by	Exhibited by
exhibits	Exhibits
expanded_form_of	Expanded form of
expected_outcome_of	Expected outcome of
finding_context_of	Finding context of
finding_site_of	Finding site of
focus_of	Focus of
form_of	Form of
has_access_instrument	Has access instrument
has_access	Has access
has_active_ingredient	Has active ingredient
has_actual_outcome	Has actual outcome
has_adjustment	Has adjustment
has_approach	Has approach
has_associated_finding	Has associated finding
has_associated_morphology	Has associated morphology
has_associated_procedure	Has associated procedure
has_branch	Has branch
has_british_form	Has British form
has_causative_agent	Has causative agent
has_challenge	Has challenge
has_component	Has component
has_conceptual_part	Has conceptual part
has_contraindicated_drug	Has contraindicated drug
has_contraindication	Has contraindication
has_course	Has course
has_definitional_manifestation	Has definitional manifestation
has_degree	Has degree
has_direct_device	Has direct device
has_direct_morphology	Has direct morphology
has_direct_procedure_site	Has direct procedure site
has_direct_substance	Has direct substance
has_divisor	Has divisor
has_dose_form	Has dose form
has_episodicity	Has episodicity
has_evaluation	Has evaluation
has_expanded_form	Has expanded form
has_expected_outcome	Has expected outcome
has_finding_context	Has finding context
has_finding_site	Has finding_site
has_focus	Has focus

has_form	Has form
has_indirect_device	Has indirect device
has_indirect_morphology	Has indirect morphology
has_indirect_procedure_site	Has indirect procedure site
has_ingredient	Has ingredient
has_intent	Has intent
has_interpretation	Has interpretation
has_laterality	Has laterality
has_location	Has location
has_manifestation	Has manifestation
has_measurement_method	Has measurement method
has_mechanism_of_action	Has mechanism of action
has_member	Has member
has_method	Has method
has_multi_level_category	Has multi-level category
has_occurrence	Has occurrence
has_onset	Has onset
has_outcome	Has associated outcome
has_part	Has part
has_pathological_process	Has pathological process
has_permuted_term	Has permuted term
has_pharmacokinetics	Has pharmacokinetics
has_physiologic_effect	Has physiologic effect
has_plain_text_form	Has plain text form
has_precise_ingredient	Has precise ingredient
has_priority	Has priority
has_procedure_context	Has procedure context
has_procedure_device	Has procedure device
has_procedure_morphology	Has procedure morphology
has_procedure_site	Has procedure site
has_process	Has process
has_property	Has property
has_recipient_category	Has recipient category
has_result	Has result
has_revision_status	Has revision status
has_scale_type	Has scale type
has_scale	Has scale
has_severity	Has severity
has_single_level_category	Has single level category
has_specimen_procedure	Has specimen procedure
has_specimen_source_identity	Has specimen source identity
has_specimen_source_morphology	Has specimen source morphology
has_specimen_source_topography	Has specimen source topography
has_specimen_substance	Has specimen substance
has_specimen	Has specimen
has_subject_relationship_context	Has subject relationship context
has_suffix	Has suffix
has_supersystem	Has supersystem
has_system	Has system
has_temporal_context	Has temporal context

has_time_aspect	Has time aspect
has_tradename	Has tradename
has_translation	Has translation
has_tributary	Has tributary
has_version	Has version
has_xml_form	Has XML form
icd_asterisk	ICD asterisk
icd_dagger	ICD dagger
indicated_by	Indicated by
indicates	Indicates
indirect_device_of	Indirect device of
indirect_morphology_of	Indirect morphology of
indirect_procedure_site_of	Indirect procedure site of
induced_by	Induced by
induces	Induces
ingredient_of	Ingredient of
intent_of	Intent of
interpretation_of	Interpretation of
interprets	Interprets
inverse_isa	Inverse is a
inverse_may_be_a	Inverse may be a
inverse_was_a	Inverse was a
is_interpreted_by	Is interpreted by
isa	Is a
larger_than	Larger Than
laterality_of	Laterality of
location_of	Location of
manifestation_of	Manifestation of
mapped_from	Mapped from
mapped_to	Mapped to
may_be_a	May be a
may_be_diagnosed_by	May be diagnosed by
may_be_prevented_by	May be prevented by
may_be_treated_by	May be treated by
may_diagnose	May diagnose
may_prevent	May prevent
may_treat	May treat
measured_by	Measured by
measurement_method_of	Measurement method of
measures	Measures
mechanism_of_action_of	Mechanism of action of
member_of_cluster	Member of cluster
metabolic_site_of	Metabolic site of
metabolized_by	Metabolized by
metabolizes	Metabolizes
method_of	Method of
modified_by	Modified by
modifies	Modifies
moved_from	Moved from
moved_to	Moved to

multiply_mapped_from	Multiply mapped from
multiply_mapped_to	Multiply mapped to
noun_form_of	Noun form of
occurs_after	Occurs after
occurs_before	Occurs before
occurs_in	Occurs in
onset_of	Onset of
other_mapped_from	Other mapped from
other_mapped_to	Other mapped to
outcome_of	Associated outcome
part_of	Part of
pathological_process_of	Pathological process of
permuted_term_of	Permuted term of
pharmacokinetics_of	Pharmakokinetics of
physiologic_effect_of	Physiologic effect of
plain_text_form_of	Plain text form of
precise_ingredient_of	Precise ingredient of
primary_mapped_from	Primary mapped from
primary_mapped_to	Primary mapped to
priority_of	Priority of
procedure_context_of	Procedure context of
procedure_device_of	Procedure device of
procedure_morphology_of	Procedure morphology of
procedure_site_of	Procedure site of
process_of	Process of
property_of	Property of
recipient_category_of	Recipient category of
replaced_by	Replaced by
replaces	Replaces
result_of	Result of
revision_status_of	Revision status of
same_as	Same as
scale_of	Scale of
scale_type_of	Scale type of
see_from	See from
see	See
severity_of	Severity of
sib_in_branch_of	Sibling in branch of
sib_in_isa	Sibling in is a
sib_in_part_of	Sibling in part of
sib_in_tributary_of	Sibling in tributary of
site_of_metabolism	Site of metabolism
smaller_than	Smaller Than
specimen_of	Specimen of
specimen_procedure_of	Specimen procedure of
specimen_source_identity_of	Specimen source identity of
specimen_source_morphology_of	Specimen source morphology of
specimen_source_topography_of	specimen source topography of
specimen_substance_of	Specimen substance of
ssc	scc

subject_relationship_context_of	Subject relationship context of
suffix_of	Suffix of
supersystem_of	Supersystem of
system_of	System of
temporal_context_of	Temporal context of
time_aspect_of	Time aspect of
tradename_of	Tradename of
translation_of	Translation of
treated_by	Treated by
treats	Treats
tributary_of	Tributary of
uniquely_mapped_from	Uniquely mapped from
uniquely_mapped_to	Uniquely mapped to
used_by	Used by
used_for	Used for
uses	Uses
use	Use
version_of	Version of
was_a	Was a
xml_form_of	XML form of
	Empty relationship attribute

STT (String Type)	
C	Varies from the preferred term only in upper-lower case
O	Other variant of the preferred form
PF	Preferred form of term
P	Plural of the preferred form
S	Singular of the preferred form
VCW	Case and word-order variant of the preferred form
VC	Case variant of the preferred form
VO	Variant of the preferred form
VW	Word-order variant of the preferred form
V	Followed by one or more of the following types of variation, in this order:
W	Contains same words as the preferred form, disregarding order and punctuation

STYPE (Column name in MRCONSO.RRF or MRREL.RRF with identifier to which attribute is attached)	
AUI	Atom identifier
CODE	Unique Identifier or code for string in source
CUI	Concept unique identifier
SAUI	Source asserted atom unique identifier
SCUI	Source asserted concept unique identifier
SDUI	Source asserted descriptor identifier
SRUI	Source asserted relationship identifier

STYPE1 (Column name in MRCONSO.RRF with first identifier to which relationship is attached)	
AUI	Atom identifier

CODE	Unique Identifier or code for string in source
CUI	Concept unique identifier
SAUI	Source asserted atom unique identifier
SCUI	Source asserted concept unique identifier
SDUI	Source asserted descriptor identifier

STYPE2	
(Column name in MRCONSO.RRF with second identifier to which relationship is attached)	
AUI	Atom identifier
CODE	Unique Identifier or code for string in source
CUI	Concept unique identifier
SAUI	Source asserted atom unique identifier
SCUI	Source asserted concept unique identifier
SDUI	Source asserted descriptor identifier

SUPPRESS	
E	Suppressible due to editor decision
N	Not suppressible
O	Obsolete, SAB,TTY may be independently suppressible
Y	Suppressible due to SAB,TTY

TOTYPE	
(Type of Expression to Which a Mapping is Mapped)	
AUI	Atom identifier
BOOLEAN_EXPRESSION	Boolean expression of strings or identifiers
CODE	Unique Identifier or code for string in source
CUI	Concept unique identifier
SAUI	Source asserted atom unique identifier
SCUI	Source asserted concept unique identifier
SDUI	Source asserted descriptor identifier
	No type, used for XR mappings.

TS (Term Status)	
P	Preferred LUI of the CUI
S	Non-Preferred LUI of the CUI
p	Preferred LUI of the CUI, suppressible (only used in ORF MRCON)
s	Non-Preferred LUI of the CUI, suppressible (only used in ORF MRCON)

TTY (Term Type in Source)	
AA	Attribute type abbreviation
AB	Abbreviation in any source vocabulary
ACR	Acronym
AC	Activities
AD	Adjective
AS	Attribute type synonym
AT	Attribute type
BD	Fully-specified drug brand name that can be prescribed

BN	Fully-specified drug brand name that can not be prescribed
BSS	Broad synonym, short
BSX	Broad synonym, expanded
BSY	Broad synonym
CCI	ISO 3166 two letter country code
CCN	Chemical code name
CCS	FIPS 10-4 country code
CCU	ISO 3166 three character country code
CC	Trimmed ICPC component process
CDA	Clinical drug name in abbreviated format
CDC	Clinical drug name in concatenated format (NDDF), Clinical drug name (NDFRT)
CDD	Clinical drug name in delimited format
CD	Clinical Drug
CE	Entry "term" to a Supplementary Concept "term"
CL	Class
CMN	Common name
CN	LOINC official component name
CO	ICPC component names (these are hierarchical terms, as opposed to the LOINC component names which are analytes)
CP	ICPC component process (in original form)
CSN	HL7 Codesystem Name
CS	Short component process in ICPC, i.e. include some abbreviations
CU	Common usage
CX	Component process in ICPC with abbreviations expanded
DE	Descriptor
DF	Dose Form
DI	Disease name
DO	Domain
DS	Short form of descriptor
DT	Definitional term, present in the Metathesaurus because of its connection to a Dorland's definition or to a definition created especially for the Metathesaurus
DX	Diagnosis
EN	MeSH nonprint entry "term"
EP	Entry "term"
EQ	Equivalent name
ESS	Exact synonym, short
ESX	Exact synonym, expanded
ESY	Exact synonym
ES	Short form of entry term
ET	Entry "term"
EX	Expanded form of entry term
FBD	Foreign brand name
FDB	Foreign drug brand
FI	Finding name
FN	Full form of descriptor
GN	Generic drug name
GO	Goal
GT	Glossary "term"
HC	Hierarchical class
HD	Hierarchical descriptor

HG	High Level Group Term
HS	Short hierarchical term (needed expansion) in ICD 10
HTN	HL7 Table Name
HT	Hierarchical term
HX	Expanded version of short hierarchical term
ID	Nursing indicator
INP	Ingredient preparation
IN	Name for an ingredient
IS	Obsolete synthesized term in the Read Thesaurus
IT	Index "term", i.e., derived from the index to any non-MeSH source vocabulary
IV	Intervention
IX	Expanded forms of indicators (embedded abbreviations expanded)
LN	LOINC official fully specified name
LO	Obsolete official fully specified name
LPDN	LOINC parts display name
LPN	LOINC parts name
LS	Expanded system/sample type (The expanded version was created for the Metathesaurus and includes the full name of some abbreviations.)
LT	Lower Level Term
LV	Lexical variant
LX	Official fully specified name with expanded abbreviations
MD	CCS multi-level diagnosis categories
MH	Main heading
MOA	Mechanism of action
MP	Preferred names of modifiers
MS	Multum names of branded and generic supplies or supplements
MT	An alternate form of a concept name from one of the source vocabularies created for the Metathesaurus
MV	Multi-level procedure category
N1	Chemical Abstracts Service Type 1 name of a chemical
NM	Supplementary chemical "term", a name of a substance
NPT	HL7 non-preferred for language term
NP	Non-preferred term
NSS	Narrow synonym, short
NSX	Narrow synonym, expanded
NSY	Narrow synonym
NS	Short form of non-preferred term
NX	Expanded form of non-preferred term
OA	Obsolete abbreviation
OBD	Obsolete branded drug
OBS	Obsolete broad synonym
OB	Obsolete term
OCD	Obsolete clinical drug
OC	Nursing outcomes
OES	Obsolete exact synonym
OF	Obsolete fully specified name
OLX	Expanded LOINC obsolete fully specified name
OL	Non-current Lower Level Term
OM	Obsolete modifiers in HCPCS
ONS	Obsolete narrow synonym

OOSN	Obsolete official short name
OPS	Obsolete preferred term
OPX	Obsolete preferred term, expanded
OP	Obsolete preferred term
OR	Orders
OSN	Official short name
OS	System-organ class in the WHO Adverse Reaction Terminology
PCE	Preferred entry "term" to a Supplementary Concept "term"
PC	Preferred "trimmed term" in ICPC
PEN	Preferred MeSH nonprint entry "term"
PEP	Preferred entry "term"
PE	Physiologic effect
PK	Pharmacokinetics
PM	Machine permutation
PN	Metathesaurus preferred name
PQ	Qualifier for a problem
PR	Name of a problem
PSC	Protocol selection criteria
PS	Short forms that needed full specification
PTGB	British preferred term
PT	Designated preferred name
PXQ	Preferred term in preferred qualifier concept.
PX	Expanded preferred terms (pair with PS)
RAB	Root abbreviation
RHT	Root hierarchical term
RLSS	Related synonym, short
RLSX	Related synonym, expanded
RLSY	Related synonym
RPT	Root preferred term
RSY	Root synonym
RS	Extracted related names in SNOMED2
RT	Designated related "term"
SA	Short forms of activities
SBDC	Semantic Branded Drug Component
SBDF	Semantic branded drug and form
SBD	Semantic branded drug
SB	Named subset of a source
SCDC	Semantic Drug Component
SCDF	Semantic clinical drug and form
SCD	Semantic Clinical Drug
SCN	Scientific name
SC	Special Category term
SD	CCS single-level diagnosis categories
SF	Synonym made by replacing ";" with no spaces around it with ", " in ICPCP2
SI	Name of a sign or symptom of a problem
SN	Official component synonym in LOINC
SP	CCS single-level procedure categories
SSN	Source short name, used in the UMLS Knowledge Source Server
SS	Synonymous "short" forms
ST	Step

SX	Mixed-case component synonym with expanded abbreviations
SYGB	British synonym
SY	Designated synonym
TA	Task
TC	Term class
TG	Name of the target of an intervention
TQ	Topical qualifier
TX	CCPSS synthesized problems for TC termgroup
UCN	Unique common name
USN	Unique scientific name
USY	Unique synonym
VAB	Versioned abbreviation
VPT	Versioned preferred term
VSY	Versioned synonym
XD	Expanded descriptor in AOD
XM	Cross mapping set
XQ	Alternate name for a qualifier
XX	Expanded string

TTY (tty_class)	
AA	abbreviation
AA	attribute
AB	abbreviation
ACR	abbreviation
AC	preferred
AD	attribute
AS	attribute
AS	synonym
AT	attribute
BD	preferred
BN	preferred
BSS	synonym
BSX	expanded
BSX	synonym
BSY	synonym
CCI	synonym
CCN	synonym
CCS	synonym
CC	preferred
CCU	synonym
CDA	abbreviation
CDC	preferred
CDD	synonym
CD	preferred
CE	entry_term
CL	hierarchical
CMN	preferred
CN	preferred
CO	hierarchical

CP	preferred
CSN	preferred
CS	abbreviation
CU	synonym
CX	expanded
CX	preferred
DE	preferred
DF	preferred
DI	preferred
DO	hierarchical
DS	abbreviation
DT	other
DX	preferred
EN	entry_term
EP	entry_term
EQ	synonym
ESS	synonym
ES	abbreviation
ES	entry_term
ESX	expanded
ESX	synonym
ESY	synonym
ET	entry_term
EX	entry_term
EX	expanded
FBD	synonym
FDB	synonym
FI	preferred
FN	preferred
GN	preferred
GO	preferred
GT	entry_term
HC	hierarchical
HD	hierarchical
HG	hierarchical
HS	abbreviation
HS	hierarchical
HTN	preferred
HT	hierarchical
HT	preferred
HX	expanded
HX	hierarchical
ID	abbreviation
ID	preferred
INP	preferred
IN	preferred
IS	obsolete
IS	synonym
IT	entry_term
IV	preferred

IX	expanded
IX	preferred
LN	preferred
LO	obsolete
LPDN	preferred
LPN	preferred
LS	expanded
LT	entry_term
LV	synonym
LX	expanded
LX	preferred
MD	preferred
MH	preferred
MOA	preferred
MP	attribute
MS	preferred
MT	preferred
MV	preferred
N1	synonym
NM	preferred
NPT	synonym
NP	synonym
NSS	synonym
NS	abbreviation
NS	synonym
NSX	expanded
NSX	synonym
NSY	synonym
NX	expanded
NX	synonym
OA	abbreviation
OA	obsolete
OBS	obsolete
OBS	synonym
OB	obsolete
OC	preferred
OES	obsolete
OES	synonym
OF	obsolete
OL	entry_term
OL	obsolete
OLX	expanded
OLX	obsolete
OLX	synonym
OM	attribute
OM	obsolete
ONS	obsolete
ONS	synonym
OOSN	obsolete
OPS	obsolete

OP	obsolete
OPX	expanded
OPX	obsolete
OR	preferred
OSN	abbreviation
OSN	preferred
OS	hierarchical
PCE	entry_term
PCE	preferred
PC	preferred
PEN	entry_term
PEN	preferred
PEP	entry_term
PEP	preferred
PE	preferred
PK	preferred
PM	synonym
PN	preferred
PQ	attribute
PR	preferred
PSC	preferred
PS	abbreviation
PT	preferred
PXQ	entry_term
PX	expanded
PX	preferred
RAB	abbreviation
RHT	hierarchical
RLSS	synonym
RLSX	expanded
RLSX	synonym
RLSY	synonym
RPT	preferred
RS	other
RSY	synonym
RT	other
SA	preferred
SBDC	preferred
SBD	preferred
SCDC	preferred
SCD	preferred
SCN	preferred
SC	other
SD	preferred
SF	synonym
SI	preferred
SN	synonym
SP	preferred
SSN	abbreviation
SS	abbreviation

SS	synonym
ST	preferred
SX	expanded
SX	synonym
SY	synonym
TA	preferred
TC	hierarchical
TG	preferred
TQ	attribute
TX	hierarchical
UCN	preferred
USN	preferred
USY	synonym
VAB	abbreviation
VPT	preferred
VSY	synonym
XD	expanded
XD	preferred
XQ	attribute
XQ	synonym
XX	expanded
XX	preferred

B.4 Source Vocabularies

All sources that contribute strings or relationships to the 2005AC Metathesaurus are listed. Each entry includes the:

- RSAB: Root Source Abbreviation
- VSAB: Versioned Source Abbreviation
- Source Official Name
- Citation: Publisher name, date and place of publication, and contact
- Number of strings included from this source

HIPAA or CHI standard vocabularies are identified.

- HIPAA standard: national standard for electronic health care transactions established by the Department of Health and Human Services under the Administrative Simplificationprovisions of the Health Insurance Portability and Accountability Act of 1996 (HIPAA, Title II)
- CHI standard: standard for use in U.S. Federal Government systems for the electronic exchange of clinical health information

Additional information on the status or use of some vocabularies in the Metathesaurus is also noted

RSAB	VSAB	Source Official Name
AIR	AIR93	AI/RHEUM, 1993

AI/RHEUM. Bethesda, (MD) : National Library of Medicine, Lister Hill Center, 1993.

> Number of Strings: 677
> Context: FULL-NOSIB-MULTIPLE

ALT
(updated)

ALT2005　　　　　　　　　　**Alternative Billing Concepts, 2005**

Alternative Billing Concepts (Altlink). Version 2005, seventh edition. Contact: Alternative Link LLC; 1065 Main St., Bldg. C, Las Cruces, NM 88005; phone: (505) 527-0636; http://www.alternativelink.com; mail@alternativelink.com.

> Number of Strings: 6892
> Context: FULL

AOD

AOD2000　　　　　　　　　　**Alcohol and Other Drug Thesaurus, 2000**

Alcohol and Other Drug Thesaurus: A Guide to Concepts and Terminology in Substance Abuse and Addiction. 3rd. ed. [4 Volumes.] Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism (NIAAA) and Center for Substance Abuse Prevention (CSAP), 2000

> Number of Strings: 20685
> Context: FULL

BI

BI98　　　　　　　　　　**Beth Israel Vocabulary, 1.0**

Beth Israel OMR Clinical Problem List Vocabulary. Version 1.0. Boston (MA): Beth Israel Deaconess Medical Center, 1999. Contact: Howard Goldberg, MD.; http://clinquery.bidmc.harvard.edu.

> Number of Strings: 1249

CCPSS

CCPSS99　　　　　　　　　　**Canonical Clinical Problem Statement System, 1999**

Canonical Clincial Problem Statement System, Version 1.0 June 23, 1999. Contact: sbrown@vumclib.mc.vanderbilt.edu

> Number of Strings: 15827

CCS
(updated)

CCS2005　　　　　　　　　　**Clinical Classifications Software, 2005**

Clinical Classifications Software (CCS). Agency for Healthcare Research and Quality (AHRQ), Rockville, MD. Release Date: April 2005. URL: http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp Phone: 301-594-1364.

> Number of Strings: 1144
> Context: FULL

CDT

CDT5　　　　　　　　　　**Current Dental Terminology 2005 (CDT-5), 5**

Current Dental Terminology 2005 (CDT 2005). Chicago, IL: American Dental Association, 2005

> Number of Strings: 590

COSTAR   **COSTAR_89-95**   **COSTAR, 1989-1995**

Computer-Stored Ambulatory Records (COSTAR). Boston (MA): Massachusetts General Hospital, 1989-1995.

> Number of Strings: 3461

CPM   **CPM2003**   **Medical Entities Dictionary, 2003**

Medical Entities Dictionary (CPM), Columbia Presbyterian Medical Center Medical Entities Dictionary. New York (NY): Columbia Presbyterian Medical Center, 2003

> Number of Strings: 3099
> Context: FULL-MULTIPLE

CPT   **CPT2005**   **Physicians' Current Procedural Terminology, 2005**

Current Procedural Terminology (CPT), 4th ed. Chicago (IL): American Medical Association, 2004. http://www.ama-assn.org

> Number of Strings: 17026
> Context: FULL-NOSIB
> HIPAA standard
> CHI standard

CPTSP   **CPT01SP**   **Physicians' Current Procedural Terminology, Spanish Translation, 2001**

Current Procedural Terminology (CPT), Spanish Translation. 4th ed. Chicago (IL): American Medical Association, 2000. http://www.ama-assn.org.

> Number of Strings: 7812

CSP
(updated)   **CSP2005**   **CRISP Thesaurus, 2005**

Computer Retrieval of Information on Scientific Projects (CRISP). Bethesda (MD): National Institutes of Health, Division of Research Grants, Research Documentation Section, 2005.

> Number of Strings: 20899
> Context: FULL-MULTIPLE

CST   **CST95**   **COSTART, 1995**

Coding Symbols for Thesaurus of Adverse Reaction Terms (COSTART). 5th ed. Rockville (MD): U.S. Food and Drug Administration, Center for Drug Evaluation and Research, 1995.

> Number of Strings: 6410
> Context: FULL-NOSIB-MULTIPLE

CTCAE
(new)   **CTCAEV3**   **Common Terminology Criteria for Adverse Events, 2003**

Common Terminology Criteria for Adverse Events v3.0, Cancer Therapy Evaluation Program (CTEP), National Cancer Institute. Bethesda, MD: National Cancer Institute, Dec. 12, 2003.

Number of Strings: 5600
Context: FULL

DDB   **DDB00**   **Diseases Database, 2000**

Diseases Database 2000. May, 2000. London (England): Medical Object Oriented Software Enterprises Ltd., 2000. Contact: Malcolm Duncan <mhduncan@compuserve.com>; http://www.diseasesdatabase.com/.

Number of Strings: 256

DMDICD10   **DMDICD10_1995**   **German translation of ICD10, 1995**

Internationale Klassifikation der Krankheiten 10 [German translation of ICD10]. Germany: Deutsches Institut fuer Medizinische Dokumentation und Information, 1998.

Number of Strings: 12002

DMDUMD   **DMDUMD_1996**   **German translation of UMDNS, 1996**

Die Nomenklatur fuer Medizinprodukte UMDNS [German translation of UMDNS]. Germany: Deutsches Institut fuer Medizinische Dokumentation und Information, 1996.

Number of Strings: 4360

DSM3R   **DSM3R_1987**   **DSM-III-R, 1987**

Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R). 3rd ed. rev. Washington (DC): American Psychiatric Association, 1987.

Number of Strings: 455
Context: FULL-NOSIB

DSM4   **DSM4_1994**   **DSM-IV, 1994**

Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). 4th ed. Washington (DC): American Psychiatric Association, 1994.

Number of Strings: 490
Context: FULL-NOSIB

DXP   **DXP94**   **DXplain, 1994**

DXplain (An expert diagnosis program). Boston (MA): Massachusetts General Hospital, 1994.

Number of Strings: 9974

GO   **GO2004_12_20**   **Gene Ontology, 2004_12_20**
(updated)

Gene Ontology. The Gene Ontology Consortium, March 2, 2004. Available from: http://www.geneontology.org/#cite_go.

Number of Strings: 25940
Context: FULL-MULTIPLE-IGNORE-RELA

HCDT **HCDT5** **HCPCS Version of Current Dental Terminology 2005 (CDT-5), 5**

HCPCS Version of Current Dental Terminology 2005 (CDT-2005). Baltimore, MD: U.S. Centers for Medicare & Medicaid, 2005.

Number of Strings: 590

HCPCS **HCPCS05** **Healthcare Common Procedure Coding System, 2005**

Healthcare Common Procedure Coding System (HCPCS). Baltimore, MD: U.S. Centers for Medicare & Medicaid Services, 2005.

Number of Strings: 5585
Context: FULL-NOSIB
HIPAA standard
CHI standard

HCPT **HCPT05** **HCPCS Version of Current Procedural Terminology (CPT), 2005**

Version of Physicians' Current Procedural Terminology (CPT) included in the Healthcare Common Procedure Coding System (HCPCS), 2005.

Number of Strings: 8665

HHC **HHC2003** **Home Health Care Classification, 2003**

Saba, Virginia. Home Health Care Classification of Nursing Diagnoses and Interventions. Washington (DC): Georgetown University, 2003.

Number of Strings: 411
Context: FULL

HL7V2.5 **HL7V2.5_2003_08_30** **HL7 Vocabulary Version 2.5, 2003_08_30**
(new)

Health Level Seven Vocabulary (HL7). Ann Arbor (MI): Health Level Seven, 1998-2002. Contact: Mark McDougall, Executive Director, Health Level Seven; 3300 Washtenaw Avenue, Suite 227, Ann Arbor, MI 48104-4250; Phone: (734)677-7777; Fax: (734)677-6622; Email: HQ@HL7.ORG ; Web site: www.HL7.ORG.

Number of Strings: 5004
Context: FULL-NOSIB

HL7V3.0 **HL7V3.0_2004_12_09** **HL7 Vocabulary Version 3.0, 2004_12_09**
(new)

Health Level Seven Vocabulary (HL7). Ann Arbor (MI): Health Level Seven, 1998-2002. Contact: Mark McDougall, Executive Director, Health Level Seven; 3300 Washtenaw Avenue, Suite 227, Ann Arbor, MI 48104-4250; Phone: (734)677-7777; Fax: (734)677-6622; Email: HQ@HL7.ORG ; Web site: www.HL7.ORG.

Number of Strings: 7135
Context: FULL-NOSIB-MULTIPLE

HLREL   **HLREL_1998**   **ICPC2E-ICD10 relationships from Dr. Henk Lamberts, 1998**

ICPC2E-ICD10 relationships from Dr. Henk Lamberts (HLREL), 1998. University of Amsterdam. Contact: H.Lamberts@AMC.UVA.NL.

HUGO   **HUGO_2005_04**   **HUGO Gene Nomenclature, 2005_04**
(updated)

HUGO Gene Nomenclature, HUGO Gene Nomenclature Committee, Department of Biology, University College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, UK. Tel: 44-20-7679-5027 Fax: 44-20-7387-3496 e-mail: nome@galton.ucl.ac.uk

   Number of Strings: 41718

ICD10   **ICD10_1998**   **ICD10, 1998**

International Statistical Classification of Diseases and Related Health Problems (ICD-10). 10th rev. Geneva (Switzerland): World Health Organization, 1998.

   Number of Strings: 13490
   Context: FULL-NOSIB

ICD10AE   **ICD10AE_1998**   **ICD10, American English Equivalents, 1998**

International Statistical Classification of Diseases and Related Health Problems (ICD-10): Americanized Version. 10th rev. Geneva (Switzerland): World Health Organization, 1998.

   Number of Strings: 1107

ICD10AM   **ICD10AM_2000**   **International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian Modification, January 2000 Release**

International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian Modification; 2nd Edition, published January 2000. Developed and Maintained by the National Centre for Classification in Health, University of Sydney, Faculty of Health Sciences. PO Box 170 Lidcombe, NSW, Australia 1825. Phone: +61 2 9351 9461. http://www.cchs.usy.edu.au/ncch/

   Number of Strings: 25588
   Context: FULL

ICD10AMAE   **ICD10AMAE_2000**   **International Statistical Classification of Diseases and Related Health Problems, Australian Modification, Americanized English Equivalents, 2000**

International Statistical Classification of Diseases and Related Health Problems, Australian Modification (ICD-10-AM), Americanized English Equivalents, produced by NLM. Bethesda (MD): National Library of Medicine, UMLS project, 2000

   Number of Strings: 2366

ICD10DUT   **ICD10DUT_200403**   **ICD10, Dutch Translation, 200403**

Hirs, W., H.W. Becker, C. van Boven, S.K. Oskam, I.M. Okkes, H. Lamberts. ICD-10, Dutch Translation, 200403. Amsterdam: Department of General Practice, Academic Medical Center/University of Amsterdam, Dutch College of General Practitioners (NHG), March 2004.

  Number of Strings: 11359

ICD9CM
(updated)   **ICD9CM_2006**   **ICD-9-CM, 2006**

ICD-9-CM [computer file]: international classification of diseases, ninth revision, clinical modification. Baltimore, MD: U.S. Department of Health and Human Services, Centers for Medicare & Medicaid Services, effective October 1, 2005.

  Number of Strings: 20396
  Context: FULL
  HIPAA standard
  CHI standard

ICPC   **ICPC93**   **International Classification of Primary Care, 1993**

The International Classification of Primary Care (ICPC). Denmark: World Organisation of Family Doctors, 1993.

  Number of Strings: 1053
  Context: FULL-NOSIB-MULTIPLE

ICPC2EDUT   **ICPC2EDUT_200203**   **International Classification of Primary Care 2nd Edition, Electronic, 2E, Dutch Translation, 200203**

Hirs, W., H.W. Becker, C. van Boven, S.K. Oskam, I.M. Okkes, H. Lamberts. International Classification of Primary Care 2E: 2nd ed. electronic. Dutch Translation. Amsterdam: Department of General Practice, Academic Medical Center/University of Amsterdam, Dutch College of General Practitioners (NHG), March 2002

  Number of Strings: 685

ICPC2EENG   **ICPC2EENG_200203**   **International Classification of Primary Care 2nd Edition, Electronic, 2E, 200203**

International Classification of Primary Care / prepared by the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/ Family Physicians (WONCA), known more briefly as the World Organization of Family Doctors. 2nd ed. Henk Lamberts and Inge Hofmans-Okkes, 2002

  Number of Strings: 749
  Context: FULL-NOSIB

ICPC2ICD10DUT
(updated)   **ICPC2ICD10DUT_200412**   **ICPC2-ICD10 Thesaurus, Dutch Translation, 200412**

International Classification of Primary Care / prepared by the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/ Family Physicians (WONCA), known more briefly as the World Organization of Family Doctors. 2nd ed. Henk Lamberts and Inge Hofmans-Okkes, Dutch, 2002

  Number of Strings: 72350

ICPC2ICD10ENG
(updated)   **ICPC2ICD10ENG_200412** **ICPC2 - ICD10 Thesaurus, 200412**

International Classification of Primary Care / prepared by the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/ Family Physicians (WONCA), known more briefly as the World Organization of Family Doctors. 2nd ed. Henk Lamberts and Inge Hofmans-Okkes, 2002

> Number of Strings: 81799

ICPC2P   **ICPC2P_2000**   **International Classification of Primary Care, Version 2-Plus, 2000**

International Classification of Primary Care, Version 2-Plus, Australian Modification. January, 2000

> Number of Strings: 13383
> Context: FULL-NOSIB-MULTIPLE

ICPCBAQ   **ICPCBAQ_1993**   **ICPC, Basque Translation, 1993**

The International Classification of Primary Care (ICPC). Basque Translation. Denmark: World Organisation of Family Doctors, 1993.

> Number of Strings: 695

ICPCDAN   **ICPCDAN_1993**   **ICPC, Danish Translation, 1993**

The International Classification of Primary Care (ICPC). Danish Translation. Denmark: World Organisation of Family Doctors, 1993.

> Number of Strings: 723

ICPCDUT   **ICPCDUT_1993**   **ICPC, Dutch Translation, 1993**

The International Classification of Primary Care (ICPC). Dutch Translation. Denmark: World Organisation of Family Doctors, 1993.

> Number of Strings: 723

ICPCFIN   **ICPCFIN_1993**   **ICPC, Finnish Translation, 1993**

The International Classification of Primary Care (ICPC). Finnish Translation. Denmark: World Organisation of Family Doctors, 1993.

> Number of Strings: 722

ICPCFRE   **ICPCFRE_1993**   **ICPC, French Translation, 1993**

The International Classification of Primary Care (ICPC). French Translation. Denmark: World Organisation of Family Doctors, 1993.

> Number of Strings: 723

ICPCGER **ICPCGER_1993** **ICPC, German Translation, 1993**

The International Classification of Primary Care (ICPC). German Translation. Denmark: World
Organisation of Family Doctors, 1993.

Number of Strings: 723

ICPCHEB **ICPCHEB_1993** **ICPC, Hebrew Translation, 1993**

The International Classification of Primary Care (ICPC). Hebrew Translation, Denmark: World
Organisation of Family Doctors, 1993

Number of Strings: 485

ICPCHUN **ICPCHUN_1993** **ICPC, Hungarian Translation, 1993**

The International Classification of Primary Care (ICPC). Hungarian Translation. Denmark: World
Organisation of Family Doctors, 1993.

Number of Strings: 718

ICPCITA **ICPCITA_1993** **ICPC, Italian Translation, 1993**

The International Classification of Primary Care (ICPC). Italian Translation. Denmark: World
Organisation of Family Doctors, 1993.

Number of Strings: 723

ICPCNOR **ICPCNOR_1993** **ICPC, Norwegian Translation, 1993**

The International Classification of Primary Care (ICPC). Norwegian Translation. Denmark: World
Organisation of Family Doctors, 1993.

Number of Strings: 722

ICPCPAE **ICPCPAE_2000** **International Classification of Primary Care, Version 2-
Plus, Americanized English Equivalents, 2000**

International Classification of Primary Care, Version 2-Plus, Australian Modification. Americanized
English Equivalents, January, 2000. Produced by NLM. Bethesda (MD): National Library of Medicine,
UMLS project

Number of Strings: 901

ICPCPOR **ICPCPOR_1993** **ICPC, Portuguese Translation, 1993**

The International Classification of Primary Care (ICPC). Portuguese Translation. Denmark: World
Organisation of Family Doctors, 1993.

Number of Strings: 723

ICPCSPA **ICPCSPA_1993** **ICPC, Spanish Translation, 1993**

The International Classification of Primary Care (ICPC). Spanish Translation. Denmark: World
Organisation of Family Doctors, 1993.

ICPCSWE   **ICPCSWE_1993**   **ICPC, Swedish Translation, 1993**

The International Classification of Primary Care (ICPC). Swedish Translation. Denmark: World Organisation of Family Doctors, 1993.

Number of Strings: 723

JABL   **JABL99**   **Online Congenital Multiple Anomaly/Mental Retardation Syndromes, 1999**

Online Congenital Multiple Anomaly/Mental Retardation Syndromes, 1999.

Number of Strings: 3257

LCH   **LCH90**   **Library of Congress Subject Headings, 1990**

Library of Congress Subject Headings. 12th ed. Washington (DC): Library of Congress, 1989.

Number of Strings: 6652

LNC   **LNC215**   **LOINC 2.15**
(updated)

Logical Observation Identifier Names and Codes (LOINC). Version 2.15. Indianapolis (IN): The Regenstrief Institute, June 6, 2005

Number of Strings: 141310
Context: FULL-NOSIB
CHI standard

MBD   **MBD05**   **MEDLINE (1995-1999)**

MEDLINE Backfiles (1995-1999). Bethesda (MD): National Library of Medicine. Contact: http://www.nlm.nih.gov.

MCM   **MCM92**   **McMaster University Epidemiology Terms, 1992**

Glossary of Methodologic Terms for Clinical Epidemiologic Studies of Human Disorders. Canada: McMaster University, 1992.

Number of Strings: 43

MDDB   **MDDB_2003_03**   **Master Drug Data Base, 2003_03**

Master Drug Data Base, 2003

Number of Strings: 9448

MDR   **MDR80**   **Medical Dictionary for Regulatory Activities**
(updated)   **Terminology (MedDRA), 8.0**

MedDRA [electronic resource] : medical dictionary for regulatory activities terminology. Version 8.0. Reston, VA : Northrop Grumman, MedDRA MSSO, September 2004.

Number of Strings: 64396
Context: FULL-MULTIPLE

MDRAE
(updated)

MDRAE80

Medical Dictionary for Regulatory Activities Terminology (MedDRA), American English Equivalents, 8.0

MedDRA : medical dictionary for regulatory activities terminology. American English Equivalents, Version 8.0. Bethesda, MD: U.S. National Library of Medicine, September 1, 2004.

Number of Strings: 659

MDRDUT

MDRDUT71

Medical Dictionary for Regulatory Activities Terminology (MedDRA), Dutch Edition, 7.1

Medical Dictionary for Regulatory Activities Terminology (MedDRA) Version 7.1, Dutch Edition. International conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Reston, VA: MedDRA MSSO, September 2004

Number of Strings: 58982
Context: FULL-MULTIPLE

MDREA
(updated)

MDREA80

Medical Dictionary for Regulatory Activities Terminology (MedDRA), American English Equivalents with expanded abbreviations, 8.0

MedDRA : medical dictionary for regulatory activities terminology, american english equivalents with expanded Abbreviations. Version 8.0. Bethesda, MD: U.S. National Library of Medicine, September 1, 2004.

Number of Strings: 16

MDREX
(updated)

MDREX80

Medical Dictionary for Regulatory Activities Terminology (MedDRA), with expanded abbreviations, 8.0

MedDRA : medical dictionary for regulatory activities terminology, with expanded abbreviations. Version 8.0. Bethesda, MD: U.S. National Library of Medicine, September 1, 2004.

Number of Strings: 472

MDRFRE

MDRFRE71

Medical Dictionary for Regulatory Activities Terminology (MedDRA), French Edition, 7.1

Medical Dictionary for Regulatory Activities Terminology (MedDRA) Version 7.1, French Edition. International conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Reston, VA: MedDRA MSSO, September 2004.

Number of Strings: 18611
Context: FULL-MULTIPLE

MDRGER
(updated)

MDRGER80

Medical Dictionary for Regulatory Activities Terminology (MedDRA), German Edition, 8.0

Medical Dictionary for Regulatory Activities Terminology (MedDRA) Version 8.0, German Edition. International conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Reston, VA: MedDRA MSSO, August 2005.

Number of Strings: 18805
Context: FULL-MULTIPLE

MDRPOR

MDRPOR71

Medical Dictionary for Regulatory Activities Terminology (MedDRA), Portuguese Edition, 7.1

Medical Dictionary for Regulatory Activities Terminology (MedDRA) Version 7.1, Portuguese Edition. International conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Reston, VA: MedDRA MSSO, September 2004.

Number of Strings: 55316
Context: FULL-MULTIPLE

MDRSPA

MDRSPA71

Medical Dictionary for Regulatory Activities Terminology (MedDRA), Spanish Edition, 7.1

Medical Dictionary for Regulatory Activities Terminology (MedDRA) Version 7.1, Spanish Edition. International conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Reston, VA: MedDRA MSSO, September 2004.

Number of Strings: 57365
Context: FULL-MULTIPLE

MED

MED05

MEDLINE (2000-2005)

MEDLINE Current Files (2000-2005). Bethesda (MD): National Library of Medicine. Contact: http://www.nlm.nih.gov.

MEDLINEPLUS

MEDLINEPLUS_20040814

MedlinePlus Health Topics_2004_08_14, 20040814

MedlinePlus Health Topics. Bethesda (MD): National Library of Medicine, August 14, 2004.

Number of Strings: 1436

MIM

MIM93

Online Mendelian Inheritance in Man, 1993

Online Mendelian Inheritance in Man (OMIM). Baltimore (MD): Johns Hopkins University, Center for Biotechnology Information, 1994.

Number of Strings: 247

MMSL

MMSL_2004_03

Multum MediSource Lexicon, 2004_03

Medisource Lexicon. Multum Information Services, Inc., Denver, CO. Release Date: March 1, 2004. URL: http://www.multum.com/ Phone: 888-633-4772 x1420.

Number of Strings: 43449

MMX	**MMX01**	**Micromedex DRUGDEX, 2001-08**

Micromedex DRUGDEX, 2001. URL: http://www.micromedex.com/ Phone: 800-525-9083.

Number of Strings: 11491

MSH **(updated)**	**MSH2006_2005_08_01**	**Medical Subject Headings, 2006_2005_08_01**

Medical Subject Headings (MeSH). Bethesda (MD): National Library of Medicine, 2006

Number of Strings: 594428
Context: FULL-MULTIPLE

MSHCZE **(updated)**	**MSHCZE2004**	**Czech translation of the Medical Subject Headings, 2004**

Czech translation of Medical Subject Headings (MeSH) 2004. Prague: Dept. of Bibliography, National Library of Medicine, 2004

Number of Strings: 26951

MSHDUT **(updated)**	**MSHDUT2005**	**Nederlandse vertaling van Mesh (Dutch translation of MeSH), 2005**

Nederlandse vertaling van MeSH [Dutch translation of MeSH), 2005. Amsterdam: Nederlands Tijdschrift voor Geneeskunde [Dutch Journal of Medicine], 2005.

Number of Strings: 41866

MSHFIN **(updated)**	**MSHFIN2005**	**Finnish translations of the Medical Subject Headings, 2005**

Finnish translations of Medical Subject Headings (MeSH), 2005. Helsinki: Finnish Medical Society Duodecim, 2005.

Number of Strings: 22582

MSHFRE	**MSHFRE2005**	**Thesaurus Biomedical Francais/Anglais [French translation of MeSH], 2005**

Thesaurus Biomedical Francais/Anglais [French translation of MeSH]. Paris (France): Institut National de la Sante et Recherche Medicale, 2005

Number of Strings: 41687
Context: FULL-NOSIB-MULTIPLE

MSHGER	**MSHGER2005**	**German translation of the Medical Subject Headings, 2005**

German translation of Medical Subject Headings (MeSH), 2005. Cologne: Deutsches Institut fur Medizinische Dokumentation und Information, 2005.

Number of Strings: 54948
Context: FULL-NOSIB-MULTIPLE

MSHITA　　　　　　　**MSHITA2005**　　　　　　**Italian translation of Medical Subject Headings, 2005**

Italian translation of Medical Subject Headings (MeSH), 2005. Rome: Istituto Superiore di Sanita, Settore Documentazione, 2005.

> Number of Strings: 23063
> Context: FULL-NOSIB-MULTIPLE

MSHJPN　　　　　　　**MSHJPN2005**　　　　　　**JAMAS Japanese Medical Thesaurus (JJMT), 2005**
(updated)

JAMAS Japanese Medical Thesaurus (JJMT). Tokyo: Japan Medical Abstracts Society; Igaku-Chuo-Zasshi, 2005.

> Number of Strings: 64341

MSHPOR　　　　　　　**MSHPOR2005**　　　　　　**Descritores em Ciencias da Saude (Portuguese translation of the Medical Subject Headings), 2005**

Descritores em Ciencias da Saude [Portuguese translation of Medical Subject Headings (MeSH)], 2004. Sao Paulo (Brazil): Latin American and Caribbean Center on Health Sciences Information. BIREME/PAHO/WHO, 2005.

> Number of Strings: 48843

MSHRUS　　　　　　　**MSHRUS2005**　　　　　　**Russian Translation of MeSH, 2005**
(updated)

Russian Translation of Medical Subject Headings (MeSH). Moscow: State Central Scientific Medical Library, 2005

> Number of Strings: 47252

MSHSPA　　　　　　　**MSHSPA2005**　　　　　　**Descritores en Ciencias de la Salud (Spanish translation of the Medical Subject Headings), 2005**

Descriptores en Ciencias de la Salud [Spanish translation of Medical Subject Headings (MeSH)], 2004. Sao Paulo: Latin American and Caribbean Center on Health Sciences Information. BIREME/PAHO/WHO, 2005.

> Number of Strings: 45326

MSHSWE　　　　　　　**MSHSWE2004**　　　　　　**Swedish translations of the Medical Subject Headings, 2004**

Swedish translation of Medical Subject Headings (MeSH), 2004. Stockholm: Karolinska Institutet, 2004.

> Number of Strings: 18143

MTH　　　　　　　　　**MTH**　　　　　　　　　　**UMLS Metathesaurus**

UMLS Metathesaurus. Bethesda, MD: National Library of Medicine.

> Number of Strings: 78145

| MTHCH | MTHCH05 | **Metathesaurus CPT Hierarchical Terms, 2005** |

Metathesaurus Hierarchical CPT Terms (these terms were created by the NLM to provide contextual information for CPT). Bethesda (MD): National Library of Medicine, 2004

Number of Strings: 938

| MTHFDA (updated) | MTHFDA_2005_03_31 | **Metathesaurus FDA National Drug Code Directory, 2005_03_31, 2005_03_31** |

Metathesaurus Forms of FDA National Drug Code Directory, 2005_03_31. Bethesda, MD: National Library of Medicine, 2005.

Number of Strings: 20817

| MTHHH | MTHHH05 | **Metathesaurus HCPCS Hierarchical Terms, 2005** |

Metathesaurus Hierarchical HCPCS Terms. Bethesda (MD): National Library of Medicine, 2005

Number of Strings: 307

| MTHHL7V2.5 (new) | MTHHL7V2.5_2003_08 | **HL7 Vocabulary Version 2.5, 7-bit equivalents, 2003_08** |

HL7 Vocabulary Version 2.5, 7-bit equivalents created by the National Library of Medicine. Bethesda (MD): August 30, 2003

Number of Strings: 2

| MTHHL7V3.0 (new) | MTHHL7V3.0_2004_12 | **HL7 Vocabulary Version 3.0, 7-bit equivalents, 2004_12** |

HL7 Vocabulary Version 3.0, 7-bit equivalents created by the National Library of Medicine. Bethesda (MD): December 9, 2004.

Number of Strings: 7

| MTHICD9 (updated) | MTHICD9_2006 | **Metathesaurus additional entry terms for ICD-9-CM, 2006** |

Metathesaurus additional entry terms for ICD-9-CM [computer file]: international classification of diseases, ninth revision, clinical modification. Bethesda, MD: U.S. Dept. of Health and Human Services, Public Health Service, National Institutes of Health, National Library of Medicine, September 2005.

Number of Strings: 18876

| MTHICPC2EAE | MTHICPC2EAE_200203 | **International Classification of Primary Care 2nd Edition, Electronic, 2E, American English Equivalents, 200203** |

Henk Lamberts and Inge Hofmans-Okkes. International Classification of Primary Care 2nd Edition, Electronic, 2E, American English Equivalents. Amsterdam: International Classification of Primary Care / prepared by the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians (WONCA), known more briefly as the World Organization of Family Doctors, 2002

MTHICPC2ICD107B   **MTHICPC2ICD107B_0412**   **ICPC2 - ICD10 Thesaurus, 7-bit Equivalents, 0412**
(updated)

International Classification of Primary Care / prepared by the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/ Family Physicians (WONCA), known more briefly as the World Organization of Family Doctors. 2nd ed. Henk Lamberts and Inge Hofmans-Okkes, 7-bit Equivalents, 2002

Number of Strings: 214

MTHICPC2ICD10AE   **MTHICPC2ICD10AE_0412**   **ICPC2 - ICD10 Thesaurus, American English**
(updated)   **Equivalents, 0412**

International Classification of Primary Care / prepared by the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/ Family Physicians (WONCA), known more briefly as the World Organization of Family Doctors. 2nd ed. Henk Lamberts and Inge Hofmans-Okkes, American English Equivalents, 2002

Number of Strings: 137

MTHMDRSPA   **MTHMDRSPA71**   **Metathesaurus Form of Medical Dictionary for**
Regulatory Activities Terminology (MedDRA), Spanish Edition, 7.1

Methathesaurus Forms of Medical Dictionary for Regulatory Activities Terminology (MedDRA) Version 7.1, Spanish Edition. Bethesda, MD: National Library of Medicine, September 2004.

Number of Strings: 27101

MTHMSHSPA   **MTHMSHSPA2005**   **Metathesaurus Form of Descritores en Ciencias de la**
(new)   **Salud (Spanish translation of the Medical Subject Headings), 2005**

Descriptores en Ciencias de la Salud [Spanish translation of Medical Subject Headings (MeSH)], 2005. Sao Paulo: Latin American and Caribbean Center on Health Sciences Information. BIREME/PAHO/ WHO, 2005.

Number of Strings: 15760

MTHMST   **MTHMST2001**   **Metathesaurus Version of Minimal Standard**
Terminology Digestive Endoscopy, 2001

Metathesaurus Version of Minimal Standard Terminology Digestive Endoscopy. Bethesda, MD: National Library of Medicine, 2001.

Number of Strings: 1944

MTHMSTFRE   **MTHMSTFRE_2001**   **Metathesaurus Version of Minimal Standard**
Terminology Digestive Endoscopy, French Translation, 2001

Metathesaurus Version of Minimal Standard Terminology Digestive Endoscopy, French Translation. Bethesda, MD: National Library of Medicine, 2001.

Number of Strings: 1833

MTHMSTITA   **MTHMSTITA_2001**   **Metathesaurus Version of Minimal Standard Terminology Digestive Endoscopy, Italian Translation, 2001**

Metathesaurus Version of Minimal Standard Terminology Digestive Endoscopy, Italian Translation. Bethesda, MD: National Library of Medicine, 2001.

Number of Strings: 1799

MTHPDQ   **MTHPDQ2005**   **Metathesaurus Forms of Physician Data Query, 2005**
(updated)

Methathesaurus Forms of Physician Data Query (PDQ), 2005. Bethesda (MD): National Cancer Institute.

Number of Strings: 49

MTHSCT   **MTHSCT_2005_07_31**   **Metathesaurus forms of SNOMED Clinical Terms, 2005_07_31**
(updated)

Metathesaurus forms of SNOMED Clinical Terms. Bethesda (MD): National Library of Medicine, July 31, 2005.

Number of Strings: 7689

MTHSCTSPA   **MTHSCTSPA_2005_04_30**   **Metathesaurus forms of Spanish SNOMED Clinical Terms, 2005_04_30**
(updated)

Metathesaurus forms of Spanish SNOMED Clinical Terms. Bethesda (MD): National Library of Medicine, April 30, 2005.

Number of Strings: 415042

NAN   **NAN2004**   **NANDA nursing diagnoses: definitions & classification, 2004**

Sparks Ralph, Sheila, Craft-Rosenberg, Martha, Herdman, T. Heather, Lavin, Mary Ann, editors. NANDA nursing diagnoses: definitions and classification 2003-2004. Philadelphia: NANDA International, 2003.

Number of Strings: 237
Context: FULL-NOSIB

NCBI   **NCBI2004_09_30**   **NCBI Taxonomy, 2004_09_30**

NCBI Taxonomy. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Library of Medicine, National Center for Biotechnology Information, September 30, 2004. http://www.ncbi.nlm.nih.gov/Taxonomy/

Number of Strings: 231959
Context: FULL-NOSIB

NCI   **NCI2004_11_17**   **NCI Thesaurus, 2004_11_17**
(updated)

NCI Thesaurus. Bethesda (MD): National Cancer Institute, National Institutes of Health, November 2004. Contact: Sherri de Coronado, decorons@exchange.nih.gov, http://nci.nih.gov

> Number of Strings: 101131
> Context: FULL-NOSIB-MULTIPLE
> HIPAA standard

NCI-CTCAE
(new)　　　　　**NCI-CTCAEV3**　　　　　**NCI modified Common Terminology Criteria for Adverse Events v3.0, 2003**

NCI modified Common Terminology Criteria for Adverse Events v3.0, Enterprise Vocabulary Service Project, National Cancer Institute. Bethesda, MD: National Cancer Institute, 2005.

> Number of Strings: 4504

NCISEER　　　　　**NCISEER_1999**　　　　　**NCI SEER ICD Neoplasm Code Mappings, 1999**

NCI Surveillance, Epidemiology, and End Results (SEER) conversions between ICD-9-CM and ICD-10 neoplasm codes. National Cancer Institute, Bethesda, MD. Release Date: June 1999. URL: http://www-seer.ims.nci.nih.gov/Admin/ConvProgs/ Phone: 301-496-8510.

NDDF
(updated)　　　　　**NDDF_2005_05_18**　　　　　**National Drug Data File Plus Source Vocabulary, May 2005, 2005_05_18, 2005_05_18**

National Drug Data File Plus Source Vocabulary 2005. San Bruno, CA: First DataBank, May 18, 2005.

> Number of Strings: 52864

NDFRT　　　　　**NDFRT_2004_01**　　　　　**National Drug File - Reference Terminology, 2004_01**

National Drug File - Reference Terminology, 2004_01. Washington, DC: U.S. Department of Veterans Affairs, Veterans Health Administration, January 2004.

> Number of Strings: 37918
> Context: FULL-NOSIB-MULTIPLE

NEU　　　　　**NEU99**　　　　　**Neuronames Brain Hierarchy, 1999**

Bowden, Douglas M., Martin, Richard F., Dubach, Joev G. Neuronames Brain Hierarchy. Seattle (WA): University of Washington, Primate Information Center, 1999. http://rprcsgi.rprc.washington.edu/neuronames/

> Number of Strings: 3864
> Context: FULL

NIC　　　　　**NIC99**　　　　　**Nursing Interventions Classification, 1999**

McCloskey, Joanne C., Bulechek, Gloria M., editors. NIC (Nursing Interventions Classification): Iowa Intervention Project. 2nd ed. St. Louis (MO): Mosby-Year Book, 1999.

> Number of Strings: 10187
> Context: FULL-NOSIB-MULTIPLE-IGNORE-RELA

NLM-MED　　　　　**NLM-MED**　　　　　**National Library of Medicine Medline Data**

National Library of Medicine (NLM) Medline Data. Bethesda (MD): National Library of Medicine. Contact: http://www.nlm.nih.gov.

NOC   **NOC97**   **Nursing Outcomes Classification, 1997**

Johnson, Marion, Maas, Meridean, editors. Nursing Outcomes Classification (NOC): Iowa Outcomes Project. St. Louis (MO): Mosby-Year Book, 1997.

> Number of Strings: 2812
> Context: FULL

OMS   **OMS94**   **Omaha System, 1994**

Martin, Karen S., Scheet, Nancy J. The Omaha System: Applications for Community Health Nursing. Philadelphia (PA): W.B. Saunders, 1992 (with 1994 corrections).

> Number of Strings: 536
> Context: FULL-MULTIPLE

PCDS   **PCDS97**   **Patient Care Data Set, 1997**

Ozbolt, Judy Grace. Patient Data Care Set (PCDS), Version 4.0, 1998.Contact: judy.ozbolt@mcmail.vanderbilt.edu; Vanderbilt University School of Nursing; 400-C Godchaux Hall; Nashville, TN 37240-0008; Telephone 615-343-3291

> Number of Strings: 2229
> Context: FULL-NOSIB

PDQ   **PDQ2005**   **Physician Data Query**
(updated)

PDQ. Bethesda (MD): National Cancer Institute, 2005.

> Number of Strings: 21974
> Context: FULL-NOSIB-MULTIPLE

PNDS   **PNDS2002**   **Perioperative Nursing Data Set, 2nd edition, 2002**
(new)

Perioperative nursing data set : the perioperative nursing vocabulary, 2nd Ed. edited by Suzanne C. Beyea. / AORN; Denver, Colo. : AORN, 2002

> Number of Strings: 268
> Context: FULL-NOSIB

PPAC   **PPAC98**   **Pharmacy Practice Activity Classification, 1998**

Pharmacy Practice Activity Classification (PPAC). Version 1. Washington (DC): American Pharmaceutical Association, 1998.

> Number of Strings: 380
> Context: FULL

PSY
(updated)

PSY2004   **Thesaurus of Psychological Index Terms, 2004**

American Psychological Association. Thesaurus of psychological index terms/Lisa A. Gallagher, ed., 10th ed., 30th anniversary, 1974-2004. Washington, DC: American Psychological Association, c2005.

Number of Strings: 7961
Context: FULL-NOSIB-MULTIPLE

QMR   **QMR96**   **Quick Medical Reference (QMR), 1996**

Quick Medical Reference (QMR). San Bruno (CA): First DataBank, 1997.

Number of Strings: 943

RAM   **RAM99**   **QMR clinically related terms from Randolph A. Miller, 1999**

QMR clinically related terms from Randolf A. Miller, 1999.

Number of Strings: 258

RCD   **RCD99**   **Clinical Terms Version 3 (CTV3) (Read Codes), 1999**

Clinical Terms Version 3 (CTV3) (Read Codes) (Q199): National Health Service National Coding and Classification Centre; March, 1999.

Number of Strings: 347308
Context: FULL-MULTIPLE

RCDAE   **RCDAE_1999**   **Read thesaurus, American English Equivalents, 1999**

American English equivalent of the Clinical Terms Version 3 (Q1, 1999), produced by NLM. Bethesda (MD): National Library of Medicine, UMLS project, 1999.

Number of Strings: 17296

RCDSA   **RCDSA_1999**   **Read thesaurus Americanized Synthesized Terms, 1999**

American English equivalent of synthesized terms from the Clinical Terms Version 3 (Q1, 1999), produced by NLM. Bethesda (MD): National Library of Medicine, UMLS project, 1999.

Number of Strings: 1180

RCDSY   **RCDSY_1999**   **Read thesaurus, Synthesized Terms, 1999**

Synthesized Read terms (without initial bracketed letters) of the Clinical Terms Version 3 (Q1, 1999), produced by NLM. Bethesda (MD): National Library of Medicine, UMLS project, 1999.

Number of Strings: 22119

RXNORM
(updated)

RXNORM_2005AC   **RXNORM Project, META2005AC Full Update 2005_09_14, 2005AC**

RxNorm work done by NLM. National Library of Medicine (NLM). Bethesda (MD): National Library of Medicine, META2005AC Full Update 2005_09_14

Number of Strings: 295134

SCTSPA
(updated)

SCTSPA_2005_04_30

SNOMED Clinical Terms, Spanish Language Edition, 2005_04_30

College of American Pathologists, SNOMED Clinical Terms, Spanish Language Edition, April 30, 2005. SNOMED International, 325 Waukegan Road, Northfield, IL 60093-2750. Phone: 800-323-4040 ext. 7700. Email: snomed@cap.org. URL: http://www.snomed.org

Number of Strings: 670485

SNM

SNM2

SNOMED-2, 2

Cote, Roger A., editor. Systematized Nomenclature of Medicine. 2nd ed. Skokie (IL): College of American Pathologists, 1979. SNOMED update, 1982. Skokie (IL): College of American Pathologists, 1982.

Number of Strings: 44062
Context: FULL-NOSIB-MULTIPLE

SNMI

SNMI98

SNOMED International, 1998

Cote, Roger A., editor. Systematized Nomenclature of Human and Veterinary Medicine: SNOMED International. Northfield (IL): College of American Pathologists; Schaumburg (IL): American Veterinary Medical Association, Version 3.5, 1998.

Number of Strings: 164173
Context: FULL-NOSIB

SNOMEDCT
(updated)

SNOMEDCT_2005_07_31

SNOMED Clinical Terms, 2005_07_31

College of American Pathologists, SNOMED Clinical Terms. SNOMED International, 325 Waukegan Road, Northfield, IL 60093-2750. Phone: 800-323-4040 ext. 7700. Email: snomed@cap.org. Release Date: July 31, 2005. URL: http://www.snomed.org

Number of Strings: 859337
Context: FULL-NOSIB-MULTIPLE
CHI standard

SPN

SPN2003

Standard Product Nomenclature, 2003

Standard Product Nomenclature (SPN). Rockville, (MD); U.S. Food and Drug Administration, 2003

Number of Strings: 4881

SRC

SRC

Metathesaurus Source Terminology Names

UMLS Metathesaurus Source Terminologies. Bethesda, MD: National Library of Medicine.

Number of Strings: 760

| ULT | ULT93 | UltraSTAR, 1993 |

Bell, Douglas. Ultrasound Structured Attribute Reporting (UltraSTAR). Boston (MA): Brigham & Womens Hospital, 1993.

Number of Strings: 84

| UMD | UMD2005 | UMDNS: product category thesaurus, 2005 |

Universal Medical Device Nomenclature System: Product Category Thesaurus. Plymouth Meeting (PA): ECRI, 2005.

Number of Strings: 17411
Context: FULL-MULTIPLE

| USPMG (new) | USPMG_2004 | USP Model Guidelines, 2004 |

United States Pharmacopeia (USP). Medicare Prescription Drug Benefit Model Guidelines: Drug Categories and Classes in Part D, 2004. http://www.usp.org/pdf/EN/mmg/comprehensiveDrugListing2004-12-31.pdf

Number of Strings: 1778
Context: FULL-MULTIPLE

| UWDA | UWDA173 | University of Washington Digital Anatomist, 1.7.3 |

University of Washington Digital Anatomist, (UWDA). Seattle (WA): University of Washinton, Version 1.7.3, March, 2003. Jose Mejino, M.D.; email: onard@biostr.washington.edu

Number of Strings: 92913
Context: FULL-MULTIPLE

| VANDF (updated) | VANDF_2005_03_23 | Veterans Health Administration National Drug File, 2005_03_23, 2005_03_23 |

U.S. Department of Veterns Affairs, Veterans Health Administration National Drug File. Department of Veterans Affairs, Washington, DC. URL: http://www.vapbm.org/PBM/natform.htm

Number of Strings: 20896

| WHO | WHO97 | WHO Adverse Reaction Terminology, 1997 |

WHO Adverse Drug Reaction Terminology (WHOART). Uppsala (Sweden): WHO Collaborating Centre for International Drug Monitoring, 1997.

Number of Strings: 3831
Context: FULL-MULTIPLE

| WHOFRE | WHOFRE_1997 | WHOART, French Translation, 1997 |

WHO Adverse Drug Reaction Terminology (WHOART). French Translation. Uppsala (Sweden): WHO Collaborating Centre for International Drug Monitoring, 1997.

Number of Strings: 3669

WHOGER   **WHOGER_1997**   **WHOART, German Translation, 1997**

WHO Adverse Drug Reaction Terminology (WHOART). German Translation. Uppsala (Sweden): WHO Collaborating Centre for International Drug Monitoring, 1997.

Number of Strings: 3401

WHOPOR   **WHOPOR_1997**   **WHOART, Portuguese Translation, 1997**

WHO Adverse Drug Reaction Terminology (WHOART). Portuguese Translation. Uppsala (Sweden): WHO Collaborating Centre for International Drug Monitoring, 1997.

Number of Strings: 3712

WHOSPA   **WHOSPA_1997**   **WHOART, Spanish Translation, 1997**

WHO Adverse Drug Reaction Terminology (WHOART). Spanish Translation. Uppsala (Sweden): WHO Collaborating Centre for International Drug Monitoring, 1997.

Number of Strings: 3104

B.5 Source and Term Type Default Order of Precedence and Suppressibility

This Appendix displays the default order of Source|Term Types and suppressibility as set by NLM and distributed in the Metathesaurus® in MRRANK.RRF or MRRANK in ORF.

Users are encouraged to change the order of Source|Term Type precedence and suppressibility to suit their requirements. The default settings will not be suitable for all applications. The highest ranking Source|Term Type within a concept determines the preferred name for that concept. Use MetamorphoSys (Section 6) to change the selection of preferred names or to alter suppressibility.

Effective with the 2005AC release, MTH|MM is no longer assigned to Metathesaurus strings with multiple meanings and has been deleted from this list. Ambiguous strings are identified in the AMBIGLUI.RRF (AMBIG.LUI in ORF) and AMBIGSUI. RRF (AMBIG.SUI in ORF) files.

Source Abbreviation	Term Type	Suppressible
MTH	PN	No
MSH	MH	No
MSH	TQ	No
MSH	PEP	No
MSH	PEN	No
MSH	EP	No
MSH	EN	No
MSH	XQ	No
MSH	PXQ	No
MSH	NM	No
RXNORM	SCD	No
RXNORM	SBD	No

RXNORM	SY	No
RXNORM	SCDF	No
RXNORM	SBDF	No
RXNORM	SCDC	No
RXNORM	DF	No
RXNORM	SBDC	No
RXNORM	BN	No
RXNORM	IN	No
DSM4	PT	No
DSM3R	PT	No
SNOMEDCT	SB	No
SNOMEDCT	XM	No
SNOMEDCT	PT	No
SNOMEDCT	FN	No
SNOMEDCT	SY	No
SNOMEDCT	PTGB	No
SNOMEDCT	SYGB	No
MTHSCT	PT	No
MTHSCT	FN	No
MTHSCT	SY	No
MTHSCT	PTGB	No
MTHSCT	SYGB	No
NCBI	SCN	No
SNMI	PT	No
SNMI	PX	Yes
SNMI	HT	No
SNMI	HX	Yes
VANDF	CD	No
VANDF	IN	No
MDDB	CD	No
USPMG	HC	No
USPMG	HT	No
USPMG	PT	No
MMX	CD	No
MMX	IN	No
RCDSA	PT	No
RCDSY	PT	No
RCDAE	PT	No
RCD	PT	No
MSH	N1	No
MSH	PCE	No
MSH	CE	No
RCDSA	OP	No
RCDSY	OP	No
RCDAE	OP	No
RCD	OP	No
SNM	PT	No

SNMI	RT	No
SNM	RT	No
SNMI	SY	No
SNMI	SX	Yes
RCDSA	SY	No
RCDSY	SY	No
RCDAE	SY	No
RCD	SY	No
RCDSA	IS	No
RCDSY	IS	No
RCDAE	IS	No
RCD	IS	No
RCDAE	AT	No
RCD	AT	No
RCD	AS	Yes
SNMI	AD	No
SNM	SY	No
SNM	RS	No
CPM	PT	No
DDB	PT	No
DDB	SY	No
NEU	HT	No
NEU	PT	No
NEU	XX	No
NEU	SY	No
UWDA	PT	No
UWDA	SY	No
UMD	PT	No
UMD	ET	No
UMD	RT	No
MMSL	CD	No
MMSL	BD	No
MMSL	SC	No
MMSL	MS	No
MMSL	GN	No
MMSL	BN	No
MMSL	IN	No
MTHFDA	CD	No
NDDF	CDC	No
NDDF	CDD	No
NDDF	CDA	No
NDDF	IN	No
NDDF	DF	No
NDFRT	CD	No
NDFRT	CDC	No
NDFRT	HT	No
NDFRT	IN	No

NDFRT	INP	No
NDFRT	DI	No
NDFRT	PE	No
NDFRT	MOA	No
NDFRT	DF	No
NDFRT	PK	No
NDFRT	SY	No
SPN	PT	No
MDRAE	HG	No
MDR	HG	No
MDREA	HG	No
MDREX	HG	No
MDRAE	PT	No
MDR	PT	No
MDREX	PT	No
MDR	OS	No
MDRAE	HT	No
MDR	HT	No
MDREA	HT	No
MDRAE	SC	No
MDREX	HT	No
MDR	SC	No
MDRAE	LT	No
MDR	LT	No
MDREA	LT	No
MDREX	LT	No
CST	PT	No
WHO	PT	No
WHO	OS	No
WHO	HT	No
WHO	IT	No
AIR	HT	No
AIR	FI	No
AIR	DI	No
AIR	SY	No
ULT	PT	No
CPT	PT	No
CPT	SY	No
CPT	MP	No
HCPT	PT	No
HCPCS	PT	No
CDT	PT	No
HCDT	PT	No
HCPCS	MP	No
HCPT	MP	No
ICD10AE	PT	No
ICD10	PT	No
ICD10AE	PX	No

ICD10	PX	No
ICD10AE	PS	Yes
ICD10	PS	Yes
ICD10AMAE	PT	No
ICD10AM	PT	No
ICD10AMAE	PX	No
ICD10AM	PX	No
ICD10AMAE	PS	Yes
ICD10AM	PS	Yes
HUGO	PT	No
HUGO	EX	No
PNDS	PX	No
PNDS	PS	No
PNDS	IV	No
PNDS	DX	No
PNDS	HT	No
PDQ	PT	No
PDQ	HT	No
PDQ	PSC	No
MTHPDQ	PT	No
PDQ	SY	No
MTHPDQ	SY	No
NCI	PT	No
NCI	HD	No
NCI	SY	No
NCI	CCI	Yes
NCI	CCN	Yes
NCI	CCS	Yes
NCI	CCU	Yes
NCI	OP	No
NCI	BN	No
NCI	FDB	No
NCI	AB	No
MTHICPC2EAE	PT	No
ICPC2EENG	PT	No
MTHICPC2ICD107B	PT	No
MTHICPC2ICD10AE	PT	No
ICPC2ICD10ENG	PT	No
ICPC	PX	No
ICPC	PT	No
ICPC	PS	Yes
ICPC	PC	No
ICPC	CX	No
ICPC	CP	No
ICPC	CS	Yes
ICPC	CC	No
ICPC2EENG	CO	No
ICPC	CO	No

MTHICPC2EAE	AB	Yes
ICPC2EENG	AB	Yes
CCPSS	TX	No
CCPSS	TC	Yes
CCPSS	PT	No
CCPSS	MP	No
ICPCPAE	SF	No
ICPCPAE	SY	No
ICPC2P	SF	No
ICPC2P	SY	No
ICPCPAE	PX	No
ICPC2P	PX	No
ICPCPAE	PT	No
ICPC2P	PT	No
ICPCPAE	PS	Yes
ICPC2P	PS	Yes
AOD	DE	No
AOD	DS	No
AOD	XD	No
AOD	FN	No
AOD	ET	No
AOD	ES	No
AOD	EX	No
AOD	NP	No
AOD	NS	No
AOD	NX	No
HCPCS	OP	No
CDT	OP	No
HCDT	OP	No
HCPT	OP	No
HCPCS	OM	No
HCPT	OM	No
GO	PX	No
GO	PT	No
GO	PS	Yes
GO	BSS	Yes
GO	BSX	No
GO	BSY	No
GO	ESS	Yes
GO	ESX	No
GO	ESY	No
GO	NSS	Yes
GO	NSX	No
GO	NSY	No
GO	OBS	Yes
GO	OES	Yes
GO	ONS	Yes
GO	RLSS	Yes

GO	RLSX	No
GO	RLSY	No
GO	SY	No
GO	OP	Yes
GO	OPX	Yes
GO	OPS	Yes
GO	IS	Yes
JABL	PC	No
JABL	PT	No
JABL	SS	No
JABL	SY	No
MIM	PT	No
PDQ	ET	No
PDQ	CU	No
PDQ	LV	No
MTHPDQ	AB	No
MTHPDQ	ACR	No
PDQ	ACR	No
PDQ	AB	No
PDQ	BD	No
MTHPDQ	FBD	No
PDQ	FBD	No
PDQ	OP	No
PDQ	IS	No
MTHPDQ	BD	No
NCBI	USN	No
NCBI	USY	No
NCBI	SY	No
NCBI	UCN	No
NCBI	CMN	No
NCBI	EQ	No
BI	PT	No
BI	SY	No
BI	RT	No
LNC	LX	No
LNC	LN	No
LNC	OSN	No
LNC	CX	No
LNC	LPN	No
LNC	LPDN	No
LNC	HC	No
LNC	HS	No
LNC	CN	No
LNC	SX	No
LNC	SN	No
LNC	XM	No
LNC	LS	No
LNC	LO	Yes

LNC	OLX	Yes
LNC	OOSN	Yes
DSM4	HT	No
DSM3R	HT	No
SNM	HT	No
ICD9CM	PT	No
MDRAE	OL	No
MDR	OL	No
MDREA	OL	No
MDREX	OL	No
ICD9CM	HT	No
CTCAE	HG	No
CTCAE	HT	No
NCI-CTCAE	HT	No
NCI-CTCAE	PT	No
CCS	HT	No
CTCAE	PT	No
CTCAE	XM	No
CCS	MD	No
CCS	SD	No
CCS	MV	No
CCS	SP	No
CCS	XM	No
ICPC2ICD10ENG	XM	No
ICD10AE	HT	No
ICD10	HT	No
ICD10AE	HX	No
ICD10	HX	No
ICD10AE	HS	Yes
ICD10	HS	Yes
ICD10AMAE	HT	No
ICD10AM	HT	No
UMD	HT	No
ICPC	HT	No
RAM	PT	No
RAM	RT	No
QMR	PT	No
MTHHL7V3.0	PT	No
HL7V3.0	PT	No
HL7V3.0	CSN	No
MTHHL7V3.0	NPT	No
HL7V3.0	NPT	No
MTHHL7V2.5	PT	No
HL7V2.5	PT	No
HL7V2.5	HTN	No
MTHCH	HT	No
MTHHH	HT	No
HHC	DX	No

BI	AB	No
HHC	IV	No
HHC	CO	No
NIC	IV	No
NIC	HC	No
NAN	PT	No
NAN	HT	No
NAN	HC	No
NAN	RT	No
OMS	MT	No
OMS	PR	No
OMS	TG	No
OMS	HT	No
OMS	PQ	No
OMS	IV	No
OMS	SI	No
NIC	AC	No
NIC	SA	No
NOC	OC	No
NOC	IX	No
NOC	ID	No
PCDS	GO	No
PCDS	OR	No
PCDS	PR	No
NIC	HT	No
NOC	HT	No
NOC	HC	No
HHC	MP	No
PCDS	CO	No
PCDS	HX	No
PCDS	HT	No
COSTAR	PT	No
DXP	DI	No
DXP	FI	No
DXP	SY	No
MCM	PT	No
MCM	RT	No
PPAC	DO	No
PPAC	CL	No
PPAC	AC	No
PPAC	ST	No
PPAC	TA	No
ALT	PT	No
ALT	SY	No
ALT	HT	No
MTH	XM	No
MTH	PT	No
MTH	SY	No

MTH	RT	No
DSM3R	SY	No
DSM3R	RT	No
MTHICD9	ET	No
CST	SC	No
CST	HT	No
CST	GT	No
PSY	PT	No
PSY	HT	No
PSY	ET	No
MTHMST	PT	No
MTHMST	SY	No
MEDLINEPLUS	PT	No
MEDLINEPLUS	XM	No
MEDLINEPLUS	ET	No
LCH	PT	No
MSH	HT	No
MSH	HS	No
MSH	PM	No
RCDSA	AB	Yes
RCDSY	AB	Yes
RCDAE	AB	Yes
RCD	AB	Yes
RCDSA	OA	Yes
RCDSY	OA	Yes
RCDAE	OA	Yes
RCD	OA	Yes
RCDAE	AA	Yes
RCD	AA	Yes
CSP	PT	No
CSP	SY	No
CSP	ET	No
CSP	AB	No
MTH	DT	No
RXNORM	ET	No
RXNORM	OCD	Yes
RXNORM	OBD	Yes
SNOMEDCT	OP	Yes
SNOMEDCT	OB	Yes
SNOMEDCT	OF	Yes
SNOMEDCT	IS	Yes
MTHSCT	OF	Yes
MTHSCT	OP	Yes
MTHSCT	IS	Yes
SCTSPA	PT	No
SCTSPA	SB	No
SCTSPA	FN	No
SCTSPA	SY	No

MTHSCTSPA	PT	No
MTHSCTSPA	FN	No
MTHSCTSPA	SY	No
MSHPOR	MH	No
MTHMSHSPA	MH	No
MSHSPA	MH	No
MSHCZE	MH	No
MSHCZE	ET	No
MSHDUT	MH	No
MSHSWE	MH	No
MSHJPN	MH	No
MSHGER	MH	No
MSHFIN	MH	No
MSHFRE	MH	No
MSHITA	MH	No
MSHITA	EP	No
MSHITA	EN	No
MSHRUS	MH	No
MSHPOR	SY	No
MTHMSHSPA	SY	No
MSHSPA	SY	No
MSHDUT	SY	No
MSHJPN	SY	No
MSHFRE	EP	No
MSHFRE	EN	No
MSHRUS	SY	No
MSHPOR	EP	No
MTHMSHSPA	EP	No
MSHSPA	EP	No
MSHGER	EP	No
MSHGER	EN	No
DMDUMD	PT	No
DMDUMD	ET	No
DMDUMD	RT	No
MDRGER	OS	No
MDRFRE	OS	No
WHOFRE	OS	No
WHOGER	OS	No
WHOPOR	OS	No
WHOSPA	OS	No
MDRSPA	HG	No
MTHMDRSPA	HG	No
MDRSPA	HT	No
MTHMDRSPA	HT	No
MDRSPA	OS	No
MTHMDRSPA	OS	No
MDRSPA	PT	No
MTHMDRSPA	PT	No

MDRSPA	SC	No
MTHMDRSPA	SC	No
MDRSPA	LT	No
MTHMDRSPA	LT	No
MDRSPA	OL	No
MDRDUT	HG	No
MDRPOR	HG	No
MDRDUT	HT	No
MDRPOR	HT	No
MDRDUT	OS	No
MDRPOR	OS	No
MDRDUT	PT	No
MDRPOR	PT	No
MDRPOR	LT	No
MDRPOR	OL	No
MDRDUT	LT	No
MDRDUT	OL	No
MDRDUT	SC	No
MDRPOR	SC	No
MDRGER	HT	No
MDRFRE	HT	No
WHOFRE	HT	No
WHOGER	HT	No
WHOPOR	HT	No
WHOSPA	HT	No
MDRGER	PT	No
MDRFRE	PT	No
MDRGER	SC	No
MDRFRE	SC	No
WHOFRE	PT	No
WHOGER	PT	No
WHOPOR	PT	No
WHOSPA	PT	No
MDRGER	HG	No
MDRFRE	HG	No
WHOFRE	IT	No
WHOGER	IT	No
WHOPOR	IT	No
WHOSPA	IT	No
CPTSP	PT	No
DMDICD10	PT	No
DMDICD10	HT	No
ICPCBAQ	PT	No
ICPCDAN	PT	No
ICPC2EDUT	PT	No
ICD10DUT	PT	No
ICD10DUT	HT	No
ICPC2ICD10DUT	PT	No

ICPCDUT	PT	No
ICPCFIN	PT	No
ICPCFRE	PT	No
ICPCGER	PT	No
ICPCHEB	PT	No
ICPCHUN	PT	No
ICPCITA	PT	No
ICPCNOR	PT	No
ICPCPOR	PT	No
ICPCSPA	PT	No
ICPCSWE	PT	No
ICPCBAQ	CP	No
ICPCDAN	CP	No
ICPCDUT	CP	No
ICPCFIN	CP	No
ICPCFRE	CP	No
ICPCGER	CP	No
ICPCHEB	CP	No
ICPCHUN	CP	No
ICPCITA	CP	No
ICPCNOR	CP	No
ICPCPOR	CP	No
ICPCSPA	CP	No
ICPCSWE	CP	No
MTHMSTFRE	PT	No
MTHMSTITA	PT	No
SRC	RPT	No
SRC	RHT	No
SRC	RAB	No
SRC	RSY	No
SRC	VPT	No
SRC	VAB	No
SRC	VSY	No
SRC	SSN	No

B.6 Release Metadata

Official Counts:
Release version: 2005AC
Release format: RRF
Concepts: 1,268,035
Number of concept names (AUIs):   5,963,467
Number of distinct concept names (SUIs):   4,945,558
Number of distinct normalized concept names (LUIs):   4,407,978
Number of sources (distinct source families by language):   118
Number of sources contributing concept names: 140
Number of languages contributing concept names: 17

Name Count by Language:

Language	Name Count	% of Metathesaurus
ENG	3965426	66.50%
SPA	1282004	21.50%
DUT	211939	3.55%
POR	133632	2.24%
GER	94241	1.58%
FRE	66578	1.12%
JPN	64341	1.08%
RUS	47252	0.79%
CZE	26955	0.45%
ITA	25586	0.43%
FIN	23304	0.39%
SWE	18866	0.32%
DAN	723	0.01%
NOR	722	0.01%
HUN	718	0.01%
BAQ	695	0.01%
HEB	485	0.01%

Name Count by Source Restriction Level (SRL):

SRL	Source Count	% of Sources
0	1914433	32.10%
1	82476	1.38%
2	22156	0.37%
3	1644701	27.58%
4	2299701	38.56%
0+4	4214134	70.66%

Count of Atoms by Suppressibility:

Suppressibility Status	Name Count	% of Metathesaurus
E	8308	0.14%
N	5418534	90.86%
O	159001	2.67%
Y	377624	6.33%

Source Counts by Language (from MRSAB):

Language	Name count	% of Metathesaurus
ENG	95	65.52%

SPA	9	6.21%
DUT	6	4.14%
GER	6	4.14%
FRE	5	3.45%
*	5	3.45%
POR	4	2.76%
ITA	3	2.07%
FIN	2	1.38%
SWE	2	1.38%
BAQ	1	0.69%
CZE	1	0.69%
DAN	1	0.69%
HEB	1	0.69%
JPN	1	0.69%
HUN	1	0.69%
NOR	1	0.69%
RUS	1	0.69%

*Note: 5 sources contribute relationships which have no associated language values.

Table of Contents

UMLS® Documentation Index

Abbreviations Used in Data Elements
Ambiguous strings
        see 2.2.7
        see 2.2.7
Ambiguous terms
AMBIG.LUI
        see 2.7.2.3.15
AMBIG.SUI
        see 2.7.2.3.16
AMBIGLUI.RRF
        see 2.7.1.3.17
AMBIGSUI.RRF
        see 2.7.1.3.18
API
ASCII
        see 6.5.2.5.1 UMLS Character Sets
Atoms
        see 2.2.3
Attribute Names
Attributes
AUI
        see 2.2.3